The Political Methodologist

Newsletter of the Political Methodology Section
American Political Science Association
Volume 17, Number 1, Fall 2009

Editors:
Paul M. Kellstedt, Texas A&M University
kellstedt@polisci.tamu.edu
David A.M. Peterson, Iowa State University
daveamp@iastate.edu
Guy D. Whitten, Texas A&M University
whitten@polisci.tamu.edu

Editorial Assistant:
Mark D. Ramirez, Washington University in Saint Louis
mdramirez@wustl.edu

Contents

Notes From the Editors

We are pleased to present you with another issue of *The Political Methodologist*. We begin this issue with some reflections on the state of political methodology by the outgoing section president, Philip Schrodt. Phil notes how the advancement of political methodology must occur side-by-side with the development of new technology. There is no doubt that technology has allowed researchers to overcome many computational problems that once seemed impossible. Many *TPM* readers find themselves estimating more complex models and having to process massive amounts of data. Thus, there is a need to find the most computationally efficient way to meet our research needs. Keeping with this theme, Dino P. Christenson and Joshua A. Morris provide both a methodology and a study of how to speed up computer processing times in R for Windows-based computers. Although a few extra seconds will be barely noticeable when dealing with simple models, their findings and suggestions should be invaluable when dealing with larger and more complex data situations.

Next, John E. Jackson provides sage advice to issues facing both junior and senior methods faculty. The issues range from getting involved in specialized conferences, overcoming the pitfalls of teaching methodology courses, and positioning one's research in regards to tenure and promotions. We would like to thank Corrine McConnaughy for putting this Q & A together.

Finally, we provide a list of announcements and awards that the section has given out over the past year. We would like to congratulate all of the winners this year and say thank you to all of those who served on the various selection committees. Finally, the president of the political methodology section, Jeff Gill, provides some opening remarks on his vision of how the section will grow under his leadership. 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

Articles

Reflections on the State of Political Methodology

Philip A. Schrodt
University of Kansas
schrodt@ku.edu

Over the past couple of years, my political methodology work has moved increasingly into the realm of applied policy analysis. As my final contribution to *TPM* as president—and irritant-in-chief—of the Society for Political Methodology, I wanted to make a three observations and four suggestions on trends I've observed in that neglected domain.

First observation: The use of what I call "technical political analysis"— systematic methodologies for forecasting, planning and policy assessment—appears to be increasing rapidly. The reasons for this would involve a separate essay, but appear to be driven by

1. the technological push of exponentially increasing computing power and the Web;
2. the generational change that has put individuals with at least some exposure to systematic political analysis—if nothing more than a single required undergraduate political science methods course—into positions of authority;
3. the widespread use of systematic methods in cognate areas such as econometric analysis and electoral forecasting.

These changes, however, appear to be occurring largely under the radar, at least as far as our graduate training is concerned.[1] The conventional wisdom remains that political methodology is primarily an esoteric ivory tower enterprise isolated from the practical world; applied analysis is dismissed in a brief and boring survey of literatures from the 1970s on quasi-experimental design and interrupted time series.

Wrong: Technical political analysis projects are now widespread, substantively and methodologically diverse, and some instances funded at multiples of the annual NSF political science program budget.[2]

Second observation: Much of this work is being directed and implemented by individuals with little or no training in political methodology. Too often—with notable exceptions—the core objectives, design and conceptualization of a project are determined before (possibly) the political scientists are brought in to work within those parameters. The results are, at best, inefficient as one spends the first part of the project cleaning up errors and re-inventions of square wheels that could easily have been avoided with a knowledge of our current research methods, and our history of looking at similar problems.[3]

Third—and in many ways, the most worrisome—observation: Many of the most ambitious and generously funded projects are being done with methodologies with which we are almost completely unfamiliar and have made little or no effort to either incorporate or refute. Most notable are the computational data-mining efforts which, with the availability of inexpensive data gathering from the Web,[4] can now easily involve gigabytes of information, with hundreds or even thousands of variables on hundreds of thousands of cases, and collected in near-real-time.

We meanwhile still spend the bulk of our time—and our graduate training—with data generated by the kindly grey-haired ladies from Michigan asking if they might please have two hours of your time to answer a series of questions about the upcoming election, or slow human coding using a meticulously constructed codebooks implemented by undergraduates making minimum wage who are more interested in each other than in either the source material or the coding protocols. In both instances, Gauss willing and the creek don't rise, maybe we will have the data assembled and ready for analysis in a year or so, and we'll get the results in the

[1] Particularly if the discussions on a certain jobs blog are any indication...

[2] Suggesting a possible new metric for funding: the "NSF", as in "Let's fund this project at about five NSFs over a three-year period." An NSF is currently about $7.5-million.

[3] The scientific press doesn't help. For example, the 24 July 2009 issue of *Science* featured a special section on complexity and networks which focused predominantly on social networks. In the introductory general-audience articles devoted largely to research on political networks, only two political scientists were quoted: the attention was on work by physicists and computer scientists. Yet in the more technical, substantive articles, about half of the authors were social scientists—economists, sociologists and political scientists—including a former president of the APSA, Elinor Ostrom. A recent widely-circulated video featuring a physicist expounding rapturously on his "discovery" of a power-law relationship known for seventy years in international relations provides another example of this problem (http://www.ted.com/talks/sean_gourley_on_the_mathematics_of_war.html).

[4] Wendy Tam Cho's comments on the "2033" panel at the Summer Meeting illustrated some of these possibilities.

APSR maybe a couple years after that. Maybe.

I was at a workshop in Europe this summer listening to the description of a very large scale computational project—in the domain of political behavior, but with only a single political scientist among the international team of thirty or more collaborators—and thought at one point, "We methodologists are a bunch of Neolithic islanders pounding rocks together to make spearheads, and just beyond the horizon sails a supercarrier with technologies we cannot even conceive of..." Consider that remark the nadir of this essay, and we will dig our way out from here.

This divergence of academic and applied methods to the detriment of the former is a relatively recent development. It was easy for the political methodology community to stay ahead of the policy community as long as the policy community did not care about systematic methodology, which pretty well describes the situation from the beginning of the Reagan administration to a year or so after 9/11/2001. However, in my world at least, this has now changed, and having regained interest, the policy community can allocate vastly more resources to analysis than we can, and they are going to rapidly outpace us if we ignore them.

But on the positive side, many of the current topics "they" are interested in are the same problems "we" are interested in. For example, a week after attending the 2009 Summer Meeting with its much-maligned emphasis on causality, I was consulting on a government project which is very interested in figuring out the state-of-the-art methodologies for evaluating... causality. Matching and rare events analysis have also featured prominently in these efforts: there is a lot of work to be done in applied contexts on hard, and very interesting, problems.

So, four suggestions.

First and foremost, I think we need to do a far better job of selling our expertise.[5] Because we do, in fact, have a great deal of expertise, both in technique and critically—as both the comments by Henry Brady and those about the late David Freedman at the Summer Meeting panels reinforced—in the areas of design, theory, conceptualization and measurement. And in the current environment, it is those aspects of political science—rather than technique alone—that provide the greatest leverage. As Larry Bartels said a number of years ago, "In a small village, the witch-doctor also needs to be a good farmer." Ultimately theory and measurement give far more leverage than technique: nothing will extract meaningful results from an incorrectly specified model estimated on bad data.

This is not to dismiss the importance of technique, but for a number of reasons—partly the increasing methodological sophistication within our discipline, but also our relatively recent ability to rapidly disseminate new methods through adaptable platforms such as Stata and R—we are now technique-rich: the diversity of methods (and applications) presented at the 2009 Summer Meeting exceeded anything I've seen in the past, and this extends a trend that has been accelerating for the last few years.[6]

We need, however, to move this expertise "upstream" in the applied world—to the specification and design of projects (and, in some instances, in determining not to waste resources on problems which simply cannot be systematically addressed). Too often, the political scientist feels like the physician facing a patient who is morbidly obese, has been smoking for thirty years and starts each day with a pint of whiskey, and says "Doc, I haven't been feeling too good lately—what can you give me?" We need some earlier interventions.

Second—reiterating a point I've been making futilely for decades—we need to actively engage the systematic alternatives to the statistical approach, notably algorithmic data-mining. And we need to engage a world where heterogeneous data flows fast and furiously—and in gigabyte quantities—via computationally-intensive Web-based extraction methods.

This actually requires two responses. The first is the methodological challenge of atheoretical data-mining, which we generally just ignore because it is, well, atheoretical data-mining, a generally highly inefficient methodology. But inefficiency is irrelevant to projects that can throw terabyte and teraflop resources at a problem.

Furthermore, data-mining clearly *does* work in *some* domains: credit card and tax fraud detection, and product recommendation systems (e.g. Amazon and Netflix) are the mostly widely known examples. Will those methods also out-perform theoretically-informed techniques in problems of systematic political analysis? I have every reason to believe that the answer is "no", but we haven't made the case. And merely dismissing, with a sneer—"That's just data-mining"—is hiding our collective heads in the sand. And won't fly with funders who are a lot more attracted to supercarriers than spearheads.

But to engage this literature we need computational skills that we have not been systematically developing. Suggesting that rather than spending the second week of graduate "math camp" on the notation for obscure metrics and line-integrals, every methodologist should know how to construct and interpret a regular expression, know the five control structures common to almost all contemporary

[5] I will return to the implications of this horrible word shortly.

[6] For example, at various points since the 1960s, one could clearly identify first contingency tables, then regression, then structural equations, then time-series and time-series cross-sections, and finally hierarchical models as the dominant focus of attention. There no longer is a dominant focus.

programming languages,[7] and be proficient in the twenty most frequently used Unix commands. Calculus was fine for Gauss and the Bernoullis, but this "Party like it's 1827!" approach to methodological pedagogy in the twenty-first century has some limitations.

Finally, we need to engage the policy community as assertive professionals based on the rich knowledge that we bring as political methodologists. We suggest revisions, we debate, we cajole, we use the best available techniques but point out their flaws as well as virtues and—critically—we turn down work when it doesn't make any sense.

Those who reject engagement with the policy community will quickly point to the risks—and they are real—of large-scale applied projects.[8] And such projects are not always easy—there have been many times when dealing with multiple layers of bureaucracies, contractors, and poorly-framed requirements that I've thought that it might be nice to just go back to, say, writing theoretical papers demonstrating that World War I couldn't occur, and therefore didn't. But systematic political analysis *is* being done, and because of the technological imperatives I noted earlier, it is going to be done whether political methodologists are involved or not. I believe that with our expertise in theory, design and method, it will be done better if our community is involved. There are risks in getting involved, but I believe the risks in not getting involved are even greater.

The question of whether any scientific community should engage in issues of practical import is not new: this issue (as well as the issues of openness and replicability) was at the core of Francis Bacon's foundational work in the early seventeenth century on what became the modern scientific approach. Of course, in the universities, Bacon's approach would continue to lose out to the established Scholastic approaches for the better part of three centuries.[9]

Several of our cognate disciplines—economics, demography and epidemiology—resolved this issue some time ago, as did the applied field of election forecasting. But for the most part we've lagged, and most (all?) of the twenty or so graduate programs capable of training (and innovating) in state-of-the-art methods seem to actively discourage students from pursuing a policy-oriented track despite the availability of jobs and funding. The prospect of publishing an incremental contribution to normal science in the *APSR* is still seen as superior than doing technical political analysis that can involve stakes of millions or even billions of dollars, and/or hundreds or even thousands of lives. I'm not convinced we serve ourselves and our communities well in that regard.

Computing and Software

A Note on Speeding Up R for Windows

Dino P. Christenson and Joshua A. Morris
The Ohio State University
christenson.24@polisci.osu.edu morris.521@polisci.osu.edu

Abstract

To what extent do different Windows PC characteristics increase the modeling efficiency of R? Do some programs or versions of R run better on different PCs? And for which kinds of models do enhanced PCs and clusters diminish processing time? This research note seeks to provide novice to intermediate level R users with a framework for understanding the benefits of explicit parallel processing and upgrades in PC hardware for large datasets and computationally burdensome models. We compare the relative benefits of each optimization with simple efficiency tests. In addition, we provide basic R code to make the transition to parallel processing easier for novice users without networked labs or cluster access.

Introduction

Today the onus of statistical modelling derives from two major sources: statistical knowledge and computational resources. This note concerns the latter. Computational resources are not limited to the operating system (OS) and

[7] except LISP

[8] Hint: before engaging in such a debate, get a promise that the first party to use the word "prostitute" has to buy a round of beer.

[9] And during the heyday of post-modernism in the 1990s, even Scholasticism was looking pretty good.

The authors' names are listed alphabetically. The authors would like to thank members of the Program in Statistics and Methodology (PRISM) and the Political Research Laboratory (PRL) at Ohio State University for encouragement and computing resources, particularly Professors Janet Box-Steffensmeier and Luke Keele as well as PRL staff Bill Miller and Isaac How; all errors are the sole responsibility of the authors. While the PRL at OSU uses Windows PCs and servers, both authors would like to note that they use Macs at home. Additional information, R code and related resources can be found at http://polisci.osu.edu/prism/resources.htm.

statistical package employed. As Sekhon (2006) notes, some operating systems are generally more efficient than others (Linux is faster than Windows XP which is faster than Mac OS X, unless the memory allocation is replaced); such is the case with statistical programs as well (Matlab is faster than R which is faster than Stata). However, substantial gains in speed can also be made within a chosen operating system and statistical package. Small upgrades in basic hardware, reformatting the data, allocating the memory and engaging explicit parallel processing all lead to relatively large jumps in the computational efficiency of statistical modelling.

We measure the extent to which hardware, package versions, cluster processing and data formatting increase the processing speed of statistical models. We focus exclusively on the interaction between the Windows XP and Vista OSs and the R statistical environment. R is increasingly popular, open source, free and, according to a recent *New York Times* article, "easy to use" for "statisticians, engineers and scientists without computer programming skills" Vance (2009). Windows, while not the statistician's current OS of choice, is the most prevalent OS in the world. Among Windows OSs, XP is the most common and Vista the latest version [1]. It is clear that regardless of the OS preferred by statisticians and programmers, the bulk of novice to intermediate applicants of statistics are still running their models on Windows. In addition, the abundance of R packages and vignettes continually make R more user-friendly to a broad spectrum of statistical modellers. Accordingly, the complex algorithms involved in various popular models can be employed without a programming background.

We believe, however, that the move to making efficiency gains in processing speed has not been dealt with as seamlessly as with the other aspects of statistical modelling. While intermediate users may make use of the models so neatly packaged in R, the benefits from such models may be mitigated by the computational burden. Should the costs be too high, students and scholars may be discouraged from using the appropriate model, in favor of less computationally intensive models and, at worst, inappropriate ones. Ultimately, we find that efficiency gains can be made through slight tweaks in the processing, hardware, version and data of models, while holding the basic OS and statistical package constant. Thus all hope for efficiency gains is not lost for Windows-based R users.

Efficiency Testing

Benchmarking processing speeds have been based on a host of models, most notably on Genetic Matching (Sekhon 2006) and bootstrapping (Rossini 2003). While the model chosen for the benchmarks will undoubtedly influence the processing speed, Sekhon (2006) recommends approaching benchmarking as a deterministic process. As such, the particular algorithm employed by the model is not critical, so long as it is consistent across machines and that one controls for the potential confounding factors. Given a lack of unobservables in computers, efficiency gains can be measured by holding all hardware and software constant and merely changing the variable of interest.

We test the processing speed with the `boot` function (Canty 1997). The `boot` function calls forth bootstrapping, a resampling method for statistical inference (Davison 1997, see also Keele 2008). While conceptually simple, bootstrapping is computationally intensive, often requiring thousands of sample replications. We follow Example 6.8 used in Davison (1997) and subsequently in Rossini's (2003) paper on simple parallel computing in R. The example seeks to demonstrate the benefit of bootstrapping on estimating the cost of constructing power plants. In the benchmarking that follows we generate bootstrap samples of a generalized linear model fit on the familiar `nuclear data` (Cox 1981). The data is comprised of the eleven variables including the dependent variable, the `cost` of the construction in millions of dollars, `date` of the construction permit and nine other seemingly relevant independent variables.

We begin by testing four different R compositions on the six different machines available in the PRL (see Table 1).[2] The different machines are identical in terms of loaded software and the version of R (2.8.0). Therefore we simply note the difference in allotted RAM, GHz, processor type, number of processors and OS, and compare the time to completion for our bootstrapping models across the machines. In each case, we run the model five times on each machine and present the average processing time across the five runs. Later we consider the potential benefit from a commercial version of R and explicit cluster processing relative to our standard R baseline. We conclude with a couple of simple tests of the benefits of data formatting and memory allocation.

[1] See W3 for a comparison of OS usage at `http://www.w3schools.com/browsers/browsers_os.asp`

[2] Note that the available machines do not allow for perfect control of the potentially confounding factors within a computer. The available machines are such that it is not possible for us to test the impact of, for example, the move from one GB of RAM to two GBs without also changing the GHz or some other feature. Thus the increase in efficiency from such a move cannot be attributed solely to the increase in GB or RAM, per se, but to an increase in both variables.

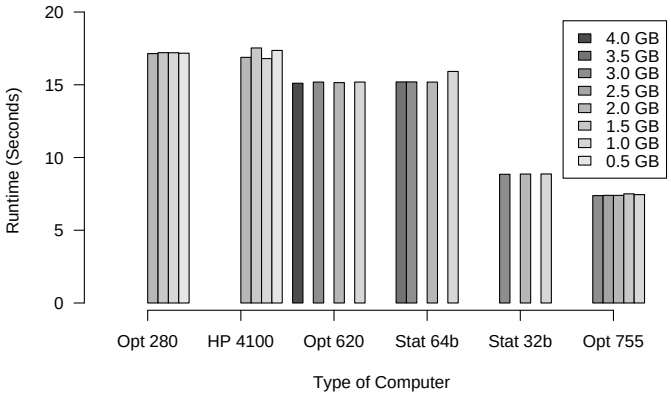
Table 1: Machines and Programs Tested

Machine	Program	RAM in GB	Processor Number	Speed in GHz	Multi Core	OS
HP 4100	*R*	.5, 1, 1.5, 2	1	2.8	no	XP
	REvolution	.5, 1, 1.5, 2	1	2.8	no	XP
	R Snow	.5, 1, 1.5, 2	1	2.8	no	XP
	ParallelR	.5, 1, 1.5, 2	1	2.8	no	XP
Optiplex 280	*R*	.5, 1, 1.5, 2	1	2.8	no	XP
	Revolution	.5	1	2.8	no	XP
	R Snow	.5	1	2.8	no	XP
	ParallelR	.5	1	2.8	no	XP
Optiplex 620	*R*	1, 2, 3, 4	1	3.2	partial	Vista
	REvolution	2	1	3.2	partial	Vista
	R Snow	2	1	3.2	partial	Vista
	ParallelR	1	1	3.2	partial	Vista
Optiplex 755	*R*	1, 1.5, 2, 2.5, 3	1	2.83	yes	Vista
	REvolution	1, 1.5, 2, 2.5, 3	1	2.83	yes	Vista
	R Snow	1, 1.5, 2, 2.5, 3	1	2.83	yes	Vista
	ParallelR	1, 1.5, 2, 2.5, 3	1	2.83	yes	Vista
32 Bit Stats	*R*	1, 2, 3	1	2.4	yes	XP
	REvolution	3	1	2.4	yes	XP
	R Snow	3	1	2.4	yes	XP
	ParallelR	2	1	2.4	yes	XP
64 Bit Stats	*R*	1, 2, 3, 3.5	2	3.2	partial	XP
	REvolution	1, 2, 3, 3.5	2	3.2	partial	XP
	R Snow	1, 2, 3, 3.5	2	3.2	partial	XP
	ParallelR	1, 2, 3, 3.5	2	3.2	partial	XP

Hardware

Upgrades in computer hardware are intended to make applications run smoother and quicker. It should be expected that hardware makes a difference in the processing speed of statistical models. We test six basic hardware factors, and hence what kind of machine composition should be used for the most efficient results, specifically: processor type, processor speed, number of processors, and the amount of RAM installed.

Figure 1: Hardware Processing Time: Memory and Speed



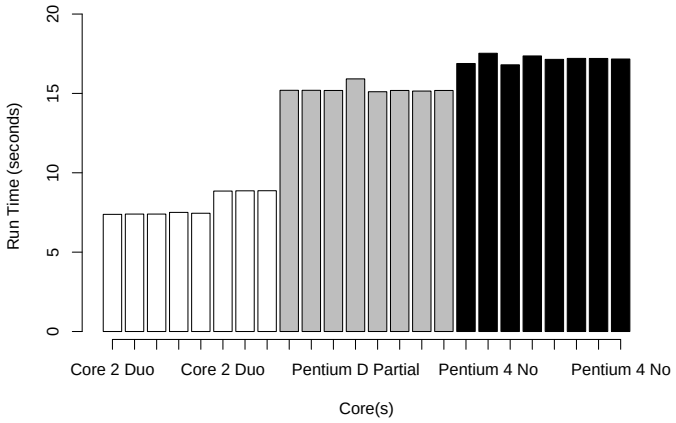
The cluster of bars denote different RAM configurations on each computer. Thus a gap between the bars are indicative of a missing RAM configuration for that computer.

Figure 1 shows that computer speed (GHz) has a minimal effect on processing time. The differences in the amount of RAM serves to increase speed for systems 2GB or greater, ortherwise it too has only a small effect on modeling efficiency. Contrarily, Figure 2 suggests that the type of processor, specifically if it is true multicore or not, is very influential. This is particularly surprising given that R does not make use of multithreaded processing (at least as of version 2.8.0). In our computation intensive test, low speed multicore processor systems perform the best. This is likely due to the fact that the processor can handle background system tasks with one core while devoting the other core to R, preventing a continual tradeoff that takes place in single core systems.

Commercial Software

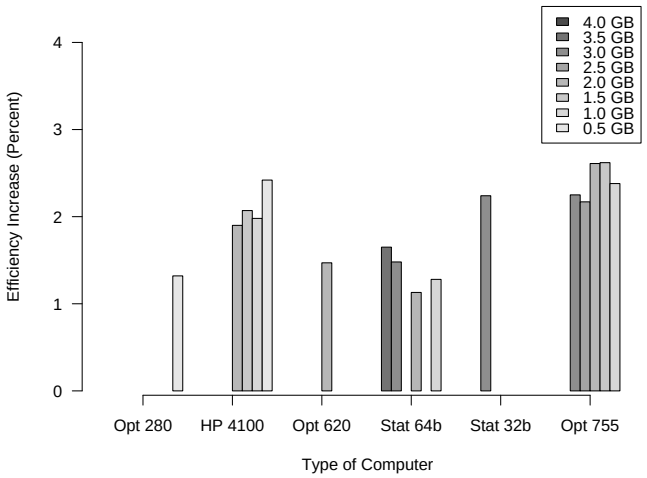
REvolution R by REvolution Computing is a commercial version of R built on the open source code base. It uses additional commercial packages and optimizations in an effort to run "many computationally-intensive programs faster." It is currently a free download with registration. We find that REvolution R outperforms the basic R in our test model across a variety of hardware. The performance increase was small, between 1.33 percent and 2.25 percent, but consistent (see Figure 3). The increase in speed was greatest for multicore processors, both greater than 2.24 percent.

Figure 2: Hardware Processing Time: Processor Core(s)



Multicore indicates two cores on a single die; the Pentium D is technically a dual core processor but the cores reside on two dies instead of one. For these tests it is considered partial multicore. The Pentium 4 No refers to a simple Pentium 4 processor without multicore. The bars denote different computer configurations.

Figure 3: Percent Increase in Efficiency of REvolution R



Parallel Processing

In computational problems the ability to process sections in parallel is highly appealing, however parallelization is not widely used. The surprising lack of use is perhaps an indicator of the troubles encountered when the statistical modeller encounters features of the programming

world. The computational difficulties encountered with parallel processing can be daunting for non-programmers.[3]. In explicit parallelization, the modeller dedicates different processes to different processors or computers. The processor which makes the calls is typically referred to as the master and the processors that carry out their given activities, the slaves. By dedicating different processes to different clusters, the statiscal program on the master node is able to take advantage of the division of labor. Each slave node is able to concentrate on its particular task and the master collects and organizes the results. In time-intensive modelling this division of labor can create enormous gains in efficiency.

We use the R package, `snow`, which appears to be the most common package for explicit parallelization and fairly user-friendly.[4] In our tests, the bootstrapping is broken up from a single set of 1000 replicates to parallel runs of 500 for a size two cluster and runs of 250 for a size four cluster. Each section of replication commands is submitted to a cluster node, where the nodes are a part of a socket cluster on the local machine. Thus we run our parallel tests on a single computer. While many advanced modellers have access to (Beowulf) clusters or networked labs, we demonstrate that efficiency gains can be made from parallel processing for novice to intermediate level users with access limited to a single machine. This means that our results can be replicated should one not have access to a networked laboratory of computers or a sophisticated cluster.[5]

Given a single machine, the process for parallelization in `snow` can be reduced to 5 simple steps. We begin by installing and loading the appropriate package.

```
# load snow library
library(snow)
```

Next, we create clusters on our computer. We utilize the `makeCluster` command to seek out nodes on the local host and specify the number of nodes to enter the cluster. Here, we also specify the type of connection between nodes: sockets, in our case, but MPI, and PVM connections are also permitted.

```
# makes a local socket cluster of size 4
c1 <- makeCluster(4, type="SOCK")
```

We can check the names and kind of processors that make up the cluster nodes with `clusterInfo` or `clusterCall`. Leaving the function expression blank returns the specified information.

```
# test cluster nodes to see name and machine
clusterCall(c1, function() Sys.info()[c("nodename",
"machine")])
```

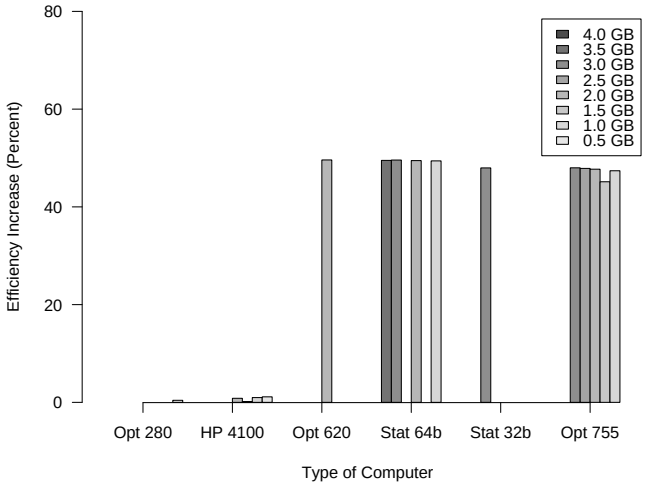
We utilize `clusterEvalQ` to evaluate a specified expression on each node. Per our example, we call it to examine the `boot` function for every node on the cluster.

```
# load boot library to cluster
clusterEvalQ(c1, library(boot))
```

We revisit the `clusterCall` command to collect the information performed at each node. `clusterCall` is the main function of the parallelization. `clusterCall` specifies arguments to be performed on each slave node and returned to the master. Finally, do not forget to call `stopCluster` at the end; otherwise, the clusters will remain connected.

```
# bootstrapping on a cluster of size 4
clusterCall(c1,boot,nuke.data,nuke.fun,R=250,m=1,
fit.pred=new.fit,x.pred=new.data))
# cleans up cluster and stops it
stopCluster(c1)
```

Figure 4: Percent Increase in Efficiency of Parallel Processing with 2 Clusters



[3]In fact, so many "user-friendly" parallel tools are in development that deciphering between them can be difficult. For example, there is an R Parellel at `http://www.rparallel.org/` that has the same objectives as REvolution's Parallel R, but is unrelated. We have not yet experimented with the former resource.

[4]R offers a host of parallel computing options. A list of R resources on parallel computing are available online at `cran.r-project.org/web/views/HighPerformanceComputing.html`. Our parallel clustering section and bootstrap model follows Rossini (2003) closely.

[5]We would expect efficiency gains from networked computers or clusters to be even greater given more time-intensive tasks.

Figure 5: Percent Increase in Efficiency of Parallel Processing with 4 Clusters

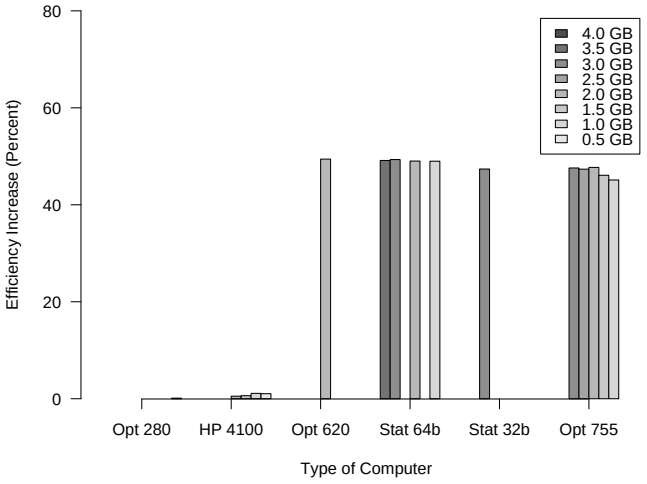
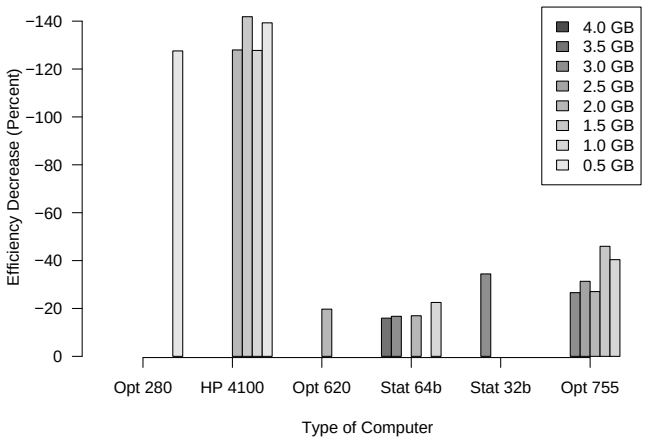


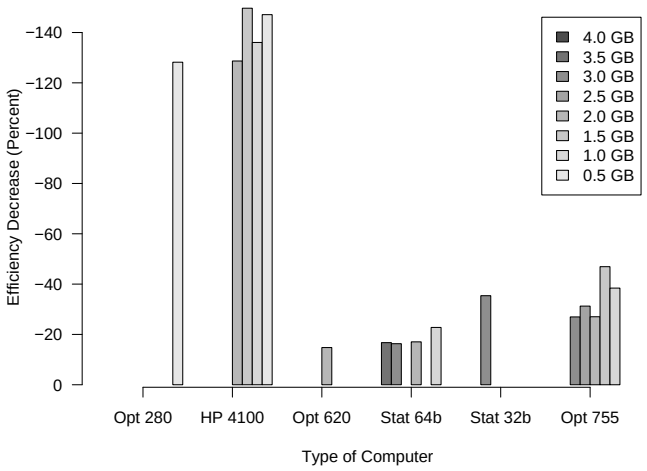
Figure 6: Percent Decrease in Efficiency of ParallelR Processing with 2 Clusters



Figures 4 and 5 show that the speed-up is substantively large in our test for machines with partial or full multicore. Multicore machines perform better than their non-clustered counterparts by over 45 percent. Single core machines see only a slight advantage, performing with less than a 1 percent speed increase over their non-clustered counterpart. We note that additional nodes do not always increase performance, as the overhead of the nodes increases with each additional node.

REvolution Computing also offers ParallelR with Enterprise, a version that includes support and automated parallel processing and optimization.[6] We compare the automated parallel processing in Enterprise with our basic R results. We found ParallelR especially user-friendly. The installation was quick and simple. The automated commands were intuitive and the move to running the new `boot` command was as simple as writing `bootNWS`. However, we do not witness any efficiency gains from the ParallelR program. Contrary to our expectations, the move to ParallelR resulted in decreases in efficiency for all machine compositions and for both size two (see Figure 6) and size four clusters (see Figure 7).[7] The results suggest that efficiency gains from parallel processing are not constant across computer compositions or types of clusters.

Figure 7: Percent Decrease in Efficiency of ParallelR Processing with 4 Clusters



Other Recommendations

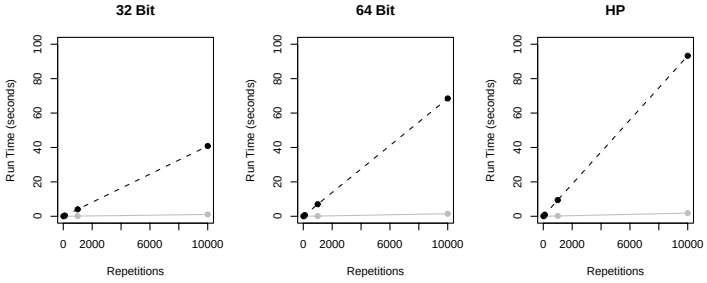
R programmers and researchers familiar with the mechanics of R have proposed additional time saving methods. Lumley (2005) of the R Core Development Team has noted various optimizations for R code, including but not limited to the following: data frames are much slower than matrices

[6] REvolution Computing was kind enough to allow us to test their ParallelR for free. More information on REvolution is available at their website `http://www.revolution-computing.com/`

[7] It is important to note, however, that we are dealing with a moderately time-intesive model, given our small dataset. In such cases, the amount of overhead it takes to supply the machine with directions for parallel processing may not be recouped by any increase in processing speed from the parallelization. In addition, it is possible that some of these programs are more efficient when run through a network cluster than on a single machine.

(especially large ones); functions that have few options and little error check are faster; and allocating memory all at once is faster than incremental allocation.

Figure 8: Processing Time of Data Frames and Matrices



We confirm these recommendations with some simple tests. We begin by running a test of the speed of data frames versus matrices. In the R environment, data frames can contain elements of mixed types while a matrix may contain elements of only one type. The test used a modified version of the `nuclear` data, called `nuclearbig`, a larger variant of `nuclear` created by duplicating the data many times over. The test consisted of comparing the differences between multiplications of the data as a matrix and as a data frame. We find that the time difference is greatest for a large number of calculations, so it is the most valuable to check what form your data is in for repeated tasks. Overall, the speed increase of using a matrix instead of a data frame in our test was around 98 percent (See Figure 8).[8]

We also checked to see if functions that have few options and little error checking are faster. Using `sum(x)/length(x)` in comparison to `mean(x)` we confirm our expectations; however, the speed difference is quite minimal for a small number of replications (see Table 2). In fact, we repeat the calculation 30,000 times before a noticeable time difference occurs. Even in this case the actual reduction was very small, less than one second. While this could result in a slight increase in speed for large repetitons, above the tens or hundreds of thousands, it is unlikely to be of substantive help otherwise. Of course, this is only one function and other function replacements may be more or less effective, but beyond this note.

We also find that memory allocation has a high overhead. We test the value of allocating memory all at once instead of incremental allocation by loading two variables full of numbers using a simple function similar to the example above. When allocating memory first using `y <- numeric(30000)` in Table 2 we find an average speed increase of 93 percent. A careful look at algorithmic approaches such as this one could make a strong difference to the run-time of the program. Of note, this optimization changes the test time from 5 seconds to 0.25 seconds on the HP 4100.

Conclusions

For the most efficiency in computationally intensive tasks in R, use multicore processor machines, if possible. Again, the gains from such a framework will not be evident on simpler models or with small to medium sized datasets, but with time-intensive models or extremely large datasets significant boosts in efficiency can be gained from utilizing the best available machines.

Programs such as REvolution R can be helpful for code that is not easily broken up for parallelization or projects which are not large enough to justify R coding tricks or rewrites; however the benefits are small relative to the other available optimizations. For computations that can be explictly parallelized, the `snow` package could be used to decrease time dramatically, even on a single machine. For bootstrapping, matching and Bayesian models, the availability of multicore processing and explicit parallel processing can be quite helpful. In these same situations it is also helpful to allocate the memory up-front. However we did not find all parallel programs to increase efficiency equally across machine compositions, and some not at all. To that end we believe that developers need to better specify the expected efficiency gains from their programs and packages, and do so with some attention to different computer compositions.

References

Canty, Angelo. 1997. "Boot Function. S original. R port" by Brian Ripley. In *Bootstrap Methods and Their Application.* Cambridge University Press.

Cox, David R. and Joyce E. Snell. 1981. *Applied Statistics: Principles and Examples.* Chapman and Hall.

Davison, Anthony C. and David V. Hinkley. 1997. *Bootstrap Methods and Their Application.* Cambridge University Press.

Keele, Luke. 2008. *Semiparametric Regression for the Social Sciences.* Wiley.

Lumley, Thomas. 2005. "R/S-PLUS Fundamentals and Programming Techniques." University of California, Berkeley.

[8] Unfortunately, this optimization is not always available, as using matrices instead of data frames is not possible for some functions, for instance, our original bootstrapping function requires data frames.

Rossini, Anthony, Luke Tierney and Na Li. 2003. "Simple Parallel Statistical Computing in R." UW Biostatistics. Working Paper Series .

Sekhon, Jasjeet Singh. 2006. "The Art of Benchmarking: Evaluating the Performance of R on Linux and OS X." *The Political Methodologist,* 14(1): 15-19.

Vance, Ashlee. 2009. "R You Ready for R?" *New York Times*.

Table 2: Average Percent Increase in Efficiency

Machine	Snow Size 2	Snow Size 4	REvolution	ParallelR Size 2	ParallelR Size 4	Matrices Frames	Less Options	Memory Up Front
Opt 755	47.22	46.77	2.41	-34.27	-34.12	96.77	88.46	91.86
32B Stats	47.97	47.38	2.24	-34.43	-35.40	94.14	86.69	92.49
64B Stats	49.49	49.12	1.38	-18.04	-18.22	97.76	88.68	89.73
Opt 620	49.60	49.42	1.47	-19.72	-14.79	98.33	88.30	91.15
HP 4100	0.77	0.83	2.09	-134.22	-140.37	97.84	88.41	96.60
Opt 280	0.42	0.12	1.32	-127.56	-128.20	98.61	89.27	90.29

Professional Development

Advice to Junior Faculty Column: Advice from John E. Jackson

John E. Jackson
University of Michigan
jjacksn@umich.edu

First, let me say these are very thoughtful and stimulating questions. I enjoyed developing answers to them, and being reminded of the pressures and opportunities I faced at the same stage. Some issues are different, but you might be a bit surprised to know how many are similar.

Expanding Networks

> Q. I really enjoy the Polmeth Summer Meeting, and am wondering about what other conferences I might attend with similar intellectual content? And what about these smaller conferences I keep hearing about? How do you find out about and get invited to those?

A. I am glad you enjoy and benefit from the Polmeth Summer Meetings. The hosts go to great length in putting on the meetings so it is nice to know their work is appreciated. The opportunities offered by the regional meetings, such as the ones organized by Jeff Gill in St. Louis and Jonathan Nagler in New York, are good examples of what you are looking for. Though they have regional names they are not restricted to scholars in those regions any more than are the Midwest meetings. These are smaller and more informal yet bring a very high quality of participant and discussion. Invitations are regularly posted on the Polmeth server.

Beyond these and the summer meetings, unfortunately, there is no simple recipe for being invited to conferences because they involve many different types, purposes, and funding arrangements. Some conferences have a broadly distributed call for papers, using venues such as the Polmeth server, *PS,* etc. Watch for these and send in a proposal if there is the slightest link from the theme to your work. At your institution become part of centers and institutes as these organizations often receive conference announcements and distribute these to faculty associated with the center. Also regularly search the websites of relevant academic and policy institutes, such as the World Bank, as they usually include calls for papers for future conferences. These conferences often have a more substantive than a methodological theme but they offer important visibility for your work and the chance to meet other scholars with interests related to

yours. In your submission make a strong connection between your research and the conference theme, even if it requires stretching one or the other a bit. Other conferences, particularly quite specialized ones, may be more selective and may have selected or announced some participants as part of the proposal/funding process. This process limits opportunities for broad access and places more weight on being part of an existing network. The most adventuresome strategy is to try to fund and organize your own conference or one done jointly with others with similar interests where you can invite other scholars you want to engage in discussion. That is precisely how the Society for Political Methodology began.

Teaching

> Q. It seems to me from conversations with other political methodologists that there's a tendency for student satisfaction with political methodology courses, especially those that are required or that satisfy some degree requirement for undergraduates, to be lower than it is with substantive courses. This seems reflected in more student complaints and lower teaching evaluations. While I don't expect my tenure outcome to rest on my teaching evaluations, I do find that the lower than average scores from my methods courses are something I am consistently asked to explain (like during my annual reviews), and that I seem to spend more time in my office with confused or complaining students than many of my colleagues. I'm beginning to feel it's just not worth the additional hassle to teach methods courses. Any advice on dealing with these issues?

A. This is a very real problem that has plagued every methodologist over the years and something we all have to find ways to overcome. Methods courses are like medicine, they are good for you but may be hard to take. There are, however, several ways to add some sugar to the medicine. Your students are Political Scientists, motivated by substantive questions. Keep the focus on interesting political questions and puzzles and teach methodological and statistical concepts by showing the insights they provide about these topics. A good insight or intuition about something the students understand and care about is worth more than a good proof. It is important they understand the formal concepts, but that is easier once they have good intuitions. A very good Polish language instructor I had said the first objective was to be able to order dinner, then he would make sure we had the grammar right. Do not, however, confuse this advice with a suggestion to teach a cookbook statistics course. It is a suggestion to use the substantive examples and applications to convey the theoretical concepts. Include in the problem sets a mix of questions. Some that require students to use and interpret the methods and others that require them to apply the concepts to situations they have not seen before.

Once the course is done and the evaluations reported are still negative, there are several things to consider. Establishing the relevant benchmark for student satisfaction in a given course is important, as is "educating" chairs and colleagues in its interpretation. After all, the design and interpretation of measures is a central part of empirical methods. Michigan's standard course evaluation form includes a question, "I was looking forward to taking this course" with responses ranging from strongly agree to strongly disagree. I have proposed that summary course evaluations should be adjusted for the responses to this question, as it is much easier to get high ratings teaching a course the students are excited about taking. This is usually not the case with a methods course, particularly if it is required.

Without such a benchmark there are several possible ways to try to create one. One benchmark is the evaluations of other required courses rather than all department courses. Possibly even disaggregating these by junior and senior faculty. If no other junior faculty teach required courses you are on your way to making your point. Some universities attempt to have evaluations comparable by summarizing evaluations only within more homogenous categories, such as all lecture courses, required versus electives, etc. A different approach is to compare your current evaluations with those from previous terms taught by someone else or even by you. Point out that your version of the course is being better received than an earlier version, if that is the case. If the improvements are relative to your earlier efforts this demonstrates improvement, which is very important. All of the techniques mentioned above take time to master and many "experiments" are required to find the most effective examples, problem sets, demonstrations and graphics. One would hope that those assessing the evaluations consider the development of effective teaching among young faculty to be as important as the evaluations themselves. Improvements speak well for the future and demonstrate you have a committment to teaching.

Research, Tenure, and Promotions

Before addressing the last three questions let me offer a general observation, which is relevant to each question in slightly different ways. This observation seconds some of Gary King's comments in the first advice column. As a junior faculty member you have, we hope, a thirty to forty year career ahead of you. Consider that faculty retiring now most likely had their first appointment in the late 1960's. Where were they then? A major task in the first five to ten years is to create the foundation that will support and sustain the scholarship and teaching you will do over that

time period. This view has several implications. Pursue the research questions and puzzles that most interest you, be they methodological, substantive, or both. It is impossible to devote the energy and time required for a successful career to research that does not stimulate your imagination and motivate you. It is a Faustian bargain to commit five years to a project merely because someone suggests it may get you tenure. Nor do you want to adopt as your primary audience the senior faculty in your current department. As strange as it may sound to you now, getting tenure at your first institution is not your primary objective. What is important is whether you are pleased with your last position, not your first or even your second. I suggest to students heading for their first job to write on an index card four institutions where you would be happy to have a tenured position in six years. Select these on the basis of the type of department and university that best fits your personality and interests—large versus small, relative time and weight given to teaching, public versus private, heterogenous or more homogenous in the type of work and approaches, etc. For the first four years look at that card at the end of each year and honestly assess whether you have made progress towards that goal. As the tenure decision gets closer look at that card more frequently to keep in mind what was your goal, how you defined success, and whether you are succeeding. If the answer to the last part is yes, then you have succeeded regardless of the current tenure decision. Now let's look at specific questions where this advice may apply.

> Q. I find myself interested in a number of methodological topics, but have some concerns about pursuing all my interests before tenure. Mostly, I'm concerned about the idea of being seen as a "jack of all trades and master of none." How much specialization within the field of methodology is necessary at this point in my career?

A. The "right" mix is what best fits your interests and work style and that applies more broadly than just the methods field. Within that framework there are some useful guidelines. Gary King, in his piece, referred to impact, which I take to mean that your work is influencing what other scholars do and is getting on their reading lists. In most cases this impact is not achieved with a single article but through an accumulation of articles and/or a book that incorporates work done post-dissertation. This requires some amount of specialization, or more accurately, a certain amount of concentration on one or two related questions. But, not too much, unless those questions are the ones that encompass your interests. Over a longer time span, your own interests as well as the value of your scholarly and teaching portfolio are very likely to be enhanced by evidence of impact on an array of topics. In fact, many departments when considering a young scholar for a tenured position want to see evidence of breadth as it suggests this person is likely to have important ideas and scholarly contributions beyond the first set of articles or book and is likely to be an active scholar in the future. A tenured position is more a statement by the tenuring institution about your future than about your past. In methods specifically it is important to demonstrate an ability to master and contribute to a number of different topics. The methods field is always evolving and it could become a problem to be typed as a specialist with only one approach.

> Q. I thought that conferences were supposed to be venues to present work in progress, so that one might benefit from and integrate ideas generated from the conversations that follow one's presentation. But I've noticed that some people—junior faculty especially— actually present work that is already under review or even accepted at a journal. I assume this is because they are trying to present work that they feel will make the best impression on the audience and improve their professional reputations. Should I be doing the same?

A. NO! I concur with, and regret, your observation about presentations at many conferences. This is shortsighted, particularly as the methods meetings provide more time to present the topic to a knowledgeable audience, for the discussant to offer comments, and to then engage the audience in conversation. Presenters do themselves and the attendees a disservice to come with a completed paper. I don't think I am the only person who is turned off by a presentation that starts by saying, "This paper provides the answer to such and such a problem, and the R-code is on my website." Meetings are not a methods bazaar where the newest products are displayed for sale.

The distinction between finished and "unfinished" work is not clear cut and will vary from paper to paper. Broadly, I take unfinished to mean on one hand that there are parts of the paper that are not fully worked out, that may still be a bit of a puzzle to the author, and/or there are still some questions about the results. Such circumstances can engage the audience and initiate a productive and possibly even unanticipated, discussion. On the other hand, enough of the paper must be completed so the discussant and audience can grasp the paper's intent, methodology, substantive application and results, and the parts that still need work. A "concept" paper or something barely past a prospectus is not a work in progress.

Obtaining the discussion and feedback that will be helpful requires that the paper be distributed to the discussant and the potential audience well in advance of the meeting. A week, or even two weeks ahead, does not provide the

discussant sufficient time to fully engage the topic, methods, proofs, data and results. There have been more than a few instances where discussants actually obtained the relevant code and/or data enabling them to replicate the paper and explore extensions and alternatives. The consequences were usually very expansive, illuminating, and productive discussions with the presenter and the audience. Far too often papers arrive, at best, the week before the conference preventing this type of interaction and engagement. At times I fear it is a bit of a defensive mechanism on the part of the authors who, incorrectly, hope to avoid this type of detailed exploration and exposure of their work.

My experience is that getting audience members engaged in and thinking about the paper and its remaining challenges provides the most stimulating sessions and solicits its valuable assistance, which may continue after the session and even after the meeting itself, with comments such as, "Have you thought about looking at the problem this way....". All papers, if they are worth anything, will lead to spirited and aggressive questioning and this may be more so with a paper that still has uncompleted issues. This, however, is where solutions to these problems, improvements, other interpretations and new ideas originate. All of which strengthen the current and future papers. The discussion may hold keys to comments reviewers may make, giving you the chance to anticipate them. This type of session may seem intimidating at the time, but a stimulating, direct and lively discussion on an important topic will be more impressive and remembered longer by attendees than a summary and discussion of a finished product about to appear in print. Finally, but not lastly, if by some chance there is a fatal flaw in the work, much better to find out at this stage and in this format than later. Trust me, I know this from experience.

At this point I want to direct a comment at our more senior colleagues, who I hope also read this column. Early in the history of the Society an important comment was made that, "We are not like the X Society. Our future is in nurturing and growing our young scholars, not in eating them alive." I believe that for the most part we were successful in establishing a norm whereby the papers, not the presenters, were rigorously questioned and dissected, which encouraged and improved both the scholarship and the presenter. This norm was particularly held to when the paper giver was a younger scholar. There were numerous instances where overly aggressive questioning of the presenter was subtly, but effectively, sanctioned. Maintaining this norm has proved to be more difficult as the Society has grown and as the collinearity of collective and personal goals has decreased. My call is to the senior members to respect and encourage the work of new scholars, but more importantly to be more vocal in responding to behavior that is not encouraging and productive for the younger scholars. Establishing a supportive environment is important for the development of all young scholars, but particularly so for our efforts to attract and retain talented young women and minority methodologists.

> Q. There seem to be "camps" within political methodology—by which I don't mean quantitative vs. qualitative, but apparent subgroups within the quantitative paradigm that might be labeled by some intellectual inclination, such as "Bayesian," "causal inference," "EITM," etc. To what extent do you think we are actually intellectually divided in this way? Are these divisions that come with consequences for publishing potential, tenure letters, and ability to get tenure at various institutions (i.e., ones where the senior faculty is of some different "camp")? How should a junior scholar navigate these divisions?

A. Camps may be a bit too strong, but certainly there are scholars of different persuasions about what statistical tools to use in different situations and about the relative benefits and costs of these tools. And as could be expected some individuals hold and express these persuasions more strongly, or narrowly, than others. It would be silly to say these persuasions do not affect article reviews, tenure letters, etc. My experience, however, in a number of roles is that these distinctions and their consequences are more muted in the methods subfield than in some other subfields. Also, keep in mind that some reviewers are likely to share your approach and demonstrate an understanding and sympathy to your work.

Good editors and departments understand and consider these persuasions when reaching decisions. This does not mean that all editors, tenure committees, chairs, deans, etc do so but many do. We all have experience with editors who simply count the number of recommendations to publish, allow one reviewer's comments to be the basis for rejection despite contrasting judgments from other reviewers, or in other ways seem to ignore the distinctions between reviewers. These experiences do not diminish with seniority, unfortunately. These are situations where it is appropriate to politely raise questions with editors. If you receive some positive reviews along with a negative one and you think the negative one has misinterpreted or misunderstood your work do two things. First read the review and your article very carefully after some time interval and explore whether your presentation may have led to the misunderstanding. Second, write the editor saying either that you now see where there was a problem or carefully outlining where and why you and the reviewer disagree and in either case would the editor consider letting you resubmit the paper. If the response is no, consider the reviews as advice, revise the paper, and

submit it to a different journal. Maybe the first journal was not the best choice for the paper in the first place. Picking the right journal for a paper is an important decision and should not be done simply on the basis of some perceived prestige or impact index but on the fit of this paper with others the journal has published and with the readership. If the issue is one of clear conflict between reviewers and the editor has sided with one reviewer, it is best to think about another journal, after carefully considering all the reviews and revising the paper. Most of the papers with which I am still most pleased went through multiple submissions and entailed discussion with editors before being accepted. I recommend this strategy only selectively, based on what you or colleagues know about the editor's style, the nature and content of the conflicting reviews and your own assessment of the paper. This brings up one last point. Do not be afraid to take a rejection to a senior colleague, either within your department or someone you have met from another department or university, and ask for advice on how to approach an editor, to think about revisions, etc. Your colleagues may not be familiar with the intricacies of the methods subfield but they certainly have had similar experiences they can draw on. Do not be afraid of "looking bad" because you got a rejection. It happens to everyone.

A forthcoming tenure review provides an opportunity to be a bit more proactive. Most reviews include a research statement, the majority of which is devoted to future plans and research topics. It is also a place to provide a roadmap for how to place your work in its relevant context. Describe how and where it fits among these different persuasions and what other authors and papers are addressed in your work, both as compliments and conflicts. If your work has become part of a dialogue among different approaches, make it clear what scholarship your work agrees with and which it does not, and why. Such statements can be folded into you future plans as you expect to further explore these debates and differences.

Thinking about this question of approaches is one place where my previous general comments apply. It is a poor idea to try to write for some particular audience or "camp" if that is not what you want your research to be. This applies both to departmental and external "audiences." That will only weaken your research in the long run. There is a standard piece of advice in sports and politics that applies here. Spend your time working on the things you can control, not on the things you cannot. Your research is your work, and the first assessment should be whether you think it is what you want it to be and is the best you can do. Those are the things you control. Chances are you will find the journal and department that share your interests and assessments of your work.

Announcements

Career Achievement Award

Nancy Burns
University of Michigan
nburns@umich.edu

We are pleased to announce the 2009 recipient of the Society for Political Methodology's Career Achievement Award. This award recognizes scholars who have made intellectual contributions that have given the field new ideas and new tools, while, at the same time, they have given the field sustaining institutions. This year's recipient is **James A. Stimson**, Raymond Dawson Bicentennial Distinguished Professor, University of North Carolina at Chapel Hill.

The citation for Jim's award that was given to the APSA reads as follows:

> Stimson led a major intellectual conceptualization of the field with his work on time-series and particularly on the analysis of pooled time-series and cross-sectional data and designs. His substantive work on issue evolution inspired his important *AJPS* paper on "Regression in Space and Time." This paper initiated a very large body of innovative methodological and applied work, much of which is still being explored. The reach of this work expands beyond Stimson's own field of American politics and is now a fixture in comparative politics and international relations, where the paper has been cited in scholarship ranging from work explaining the number of parties in Argentina to work exploring the determinants of international trade. His work with aggregate time-series data stimulated many important methodological and substantive discussions and papers and was one of the first uses

of Box-Jenkins time series methods in Political Science. Stimson originated and provided definition and direction for the use and understanding of these methods in the field.

His work with Edward Carmines on issue evolution and the long-term connection between parties, the mass public, and representation has had a tremendous impact. His solo work on the nature of public opinion and public policy mood reshaped how scholars think about public opinion. His collaboration with Michael MacKuen and Robert Erikson on The Macro Polity challenged the conventional wisdom on partisanship and extant understandings about the link between economics and politics. His work on public mood led him to create the time series measurement algorithm CALC which has been used by numerous other scholars for their own applications. For The Macro Polity, Stimson and his collaborators were early pioneers in work with the DYMIMIC estimator to model the dynamic link between time series with multiple indicators. On the measurement side, his public mood scale is the most widely used measure of public liberalism across time at the macro-level.

Stimson's work has been widely recognized and has received numerous prestigious awards. His book, *Issue Evolution*, with Carmines received the Kammerer Award in 1990 as the APSA's best book in American politics; *Tides of Consent* received the 2006 Goldsmith Prize from the Shorenstein Center at the John F. Kennedy School for the best book on politics, the press and public affairs; in 1996 he shared the Heinz Eulau prize for the best paper published in the *APSR* the previous year; and in 2005 he shared the McGraw-Hill Award for the best paper published on law and courts. In 2000 he was elected to the American Academy of Arts and Sciences. His scholarly work is widely praised, and the breadth of topics is impressive.

Stimson has served the Society in almost every way possible. Jim attended the first Summer Political Methodology Workshop in Ann Arbor in July, 1984. This workshop laid the foundation for the Society for Political Methodology and the now twenty-six year long series of summer conferences that have grown from fifteen to three hundred participants. He served as the organization's president from 1995-1997.

Stimson is also responsible for one of the Society's most important institutions. He was the original editor of our very successful journal, *Political Analysis*. His work to establish *Political Analysis* as a major journal at a time when the organization barely existed and then his editorial leadership for the first three issues created the journal we now have and value. His vision for the journal and his incredible energy, patience, and persistence are evident in the journal's reputation and impact.

Finally, Stimson has been a tremendously successful mentor and collaborator in the field.

This year's award committee consisted of:
Nancy Burns, University of Michigan (Chair); Jake Bowers, University of Illinois, Urbana-Champaign; Janet Box-Steffensmeier, The Ohio State University; John Jackson, University of Michigan; and Tse-Min Lin, University of Texas, Austin.

Fellows of the Society for Political Methodology

John R. Freeman
University of Minnesota
freeman@umn.edu

Congratulations to the 2009 Fellows of the Society for Political Methodology. This award acknowledges individuals who have made outstanding scholarly contributions to the field of political methodology. Specifically, "selection to the position of Fellow of the Society for Political Methodology honors individuals who have made outstanding scholarly contributions to the development of political methodology, and whose methodological work has had a major international impact on subsequent scholarship in the field, in the discipline more broadly, and where appropriate in other areas."

Charles H. Franklin received his B.A. from Birmingham-Southern College and his Ph.d. from the University of Michigan. He has held permanent positions at Washington University in St. Louis and at the University of Wisconsin-Madison. For many years he has taught courses at the ICPSR and the Essex Summer School in Social Science Data Analysis and Collection. In addition, for eight years he served on the Board of Overseers for the National Election Study. He has been a consultant to ABC News and to the Voter News Service. Charles also is a co-founder of the important website Pollster.com. He has (co-)authored numerous articles on a broad array of topics in Political Methodology and American politics including socialization, voter behavior, electoral politics, and judicial politics. His paper, "Estimation Across Datasets: Two-stage Auxiliary Instrumental Variables Estimation" (*Politi-*

cal Analysis, 1990), received honorable mention as the best paper presented at the 1988 Annual Meeting of the American Political Science Association. His co-authored article with Liane Kosaki, "The Republican Schoolmaster: The Supreme Court, Public Opinion and Abortion," (*American Political Science Review*, 1989) has been reprinted several times; this article recently received the 2009 Lasting Contribution Award of the A.P.S.A.'s Law and Courts Section. Charles was President of the Society for Political Methodology from 1999-2001. He hosted the summer methods conference at two different institutions.

Keith T. Poole currently is professor of political science at the University of California, San Diego. He received his undergraduate degree from Portland State University and his M.A. and Ph.d. from the University of Rochester. Keith is a member of the American Academy of Arts and Sciences; he has been a fellow at the Center for Advanced Studies in the Behavioral Sciences. Among his many honors are the Warren E. Miller Award for the best article in the 2001 volume of *Political Analysis* and the Duncan Black Prize for the best article by a senior scholar in the 2007 issue of *Public Choice*. The (co)author of five books, an important monograph, numerous scholarly articles, and theoretically important data sets and software. Keith is perhaps best well known nationally and internationally for his work with Howard Rosenthal in spatial modeling, particularly, his and Rosenthals models of legislatures. Illustrative of his contributions to this field is Keith's 2005 book *Spatial Models of Parliamentary Voting* (Cambridge University Press, 2005), a book that is rapidly becoming a classic. The National Science Foundation supported his research several times. He also won support from the Carnegie-Bosch Foundation.

Howard Rosenthal received both his B.S. and Ph.d degrees from the Massachusetts Institute of Technology. He currently is Professor of Politics at New York University. Howard is a member of the American Academy of Arts and Sciences. Twice he has been a fellow at the Center for Advanced Study in Behavioral Sciences. He also has held fellowships at research institutes in several European countries and in Israel. Among his many awards is the first Statistical Software Award given by the Political Methodology Society (with Keith Poole) for NOMINATE. Seminal contributions to economics, political science and sociology have been made by Howard. His work with Keith Poole on spatial modeling of legislatures is recognized as one of the most important in political science in recent decades. An example of their work is the classic: *Congress: A Political-Economic History of Roll Call Voting* (Oxford University Press 1997). In addition to his work with Poole, he has had three other important collaborations. In work, with Alberto Alesina, Howard has shown how partisan politics and other features of democratic politics are a cause and a consequence of certain market processes (*Partisan Politics, Divided Government and the Economy* Cambridge University Press 1997). In work with Thomas Palfrey, he advanced our understanding in voluntary participation in voting and in the provision of public goods. In work with Thomas Romer, he developed the agenda setter model that has influenced the institutional study of Congress. In so doing, Howard has been a leader in building bridges between the theoretical and empirical branches of political economy as advocated by the proponents of EITM. Many foundations have supported his research (multiple times) including the National Science Foundation, Russell Sage Foundation, Spencer Foundation and Ford Foundation.

Gosnell Award

Kenneth Kollman
University of Michigan
kkollman@umich.edu

This year's Gosnell Award winners for best research presented on political methodology at a conference are **John Freeman**, of the University of Minnesota, and **Jeff Gill**, of Washington University in Saint Louis. Their paper, "Dynamic Elicited Priors for Updating Covert Networks" was presented at the Political Methodology annual meetings in July, 2008.

The Freeman and Gill research develops a Bayesian method to combine elicited priors from experts to improve inference about the nature of social networks. The application in the paper is on networks where individuals have an interest in keeping their connections hidden. Terrorist networks are examples. One of the real strengths of the research lies in the prospects for extensions to problems across political science. The paper improves on our methods for collecting and analyzing hard-to-quantify information about organizations, networks, and relationships.

John Williams Dissertation Award

Patrick Brandt
University of Texas, Dallas
pbrandt@utdallas.edu

The John Williams Dissertation Proposal Award committee (Patrick Brandt [chair], Michael Colaresi, and Betsy Sinclair) has selected **Xun Pang's** dissertation proposal as this year's winner.

The proposal, "A Bayesian Probit Hierarchical Model with AR(p) Errors and Non-nested Clustering:

Studying Sovereign Creditworthiness and Political Institutions", outlines the implementation of a Bayesian generalized linear multilevel model with pth-order autoregressive errors for modeling unbalanced binary and ordinal TSCS data. The committee believes that the resulting multilevel models will be widely applicable to many kinds of longitudinal data analyses in political science and political economy by providing ways to model both the dynamics and heterogeneity of limited dependent variable data. The proposed implementation will also include software and new estimation methods that should have impact as well.

Congratulations to Xun Pang on this accomplishment!

Statistical Software Award

Jasjeet S. Sekhon
University of California, Berkeley
sekhon@berkeley.edu

On behalf of the Political Methodology section, we congratulate **Keith Poole** and **Howard Rosenthal** on winning our first annual Statistical Software Award for their work on NOMINATE. The award committee—Micah Altman, Kosuke Imai, Andrew Martin, Simon Jackman and Jasjeet Sekhon (chair)—thank all who submitted nominations. The award recognizes individuals for developing statistical software that makes a significant research contribution. We also judge the software's source code for quality and algorithmic innovation.

NOMINATE is a landmark in software development for political science. It allowed researchers to extract from roll call votes the small number of fundamental dimensions that generate the issue specific dimensions. NOMINATE continues to allow researchers to analyze roll call votes consistent with the spatial theory of voting. The software has been enormously influential: estimates produced by it have been used in hundreds of books and articles and publications that rely on NOMINATE have been cited thousands of times.

A Note from our Section President

Jeff Gill
Washington University in Saint Louis
jgill@wustl.edu

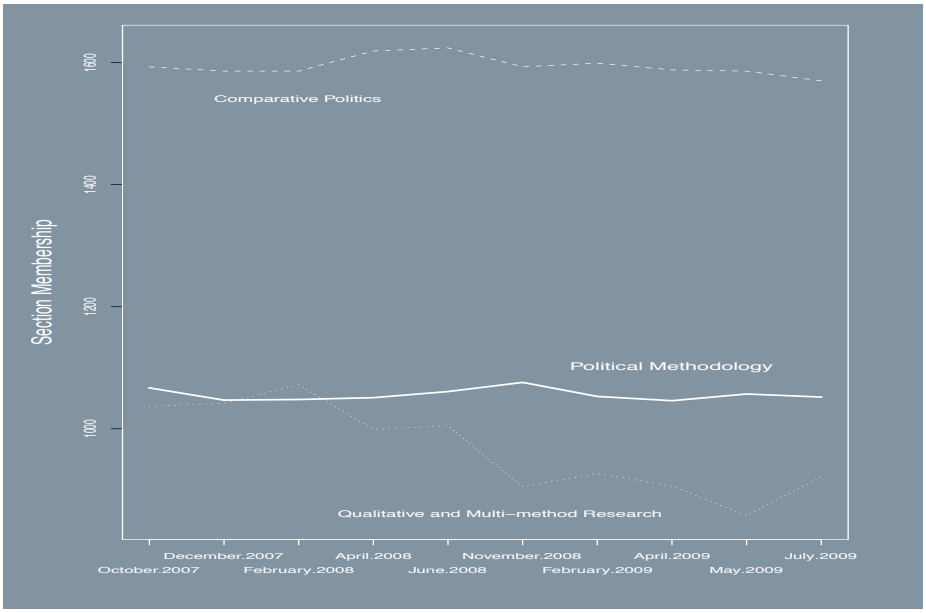
I'm honored and humbled to be writing in the capacity of President of the Society for Political Methodology and the APSA Section on Political Methodology. We have the top journal in the field for two years in a row, our Summer Meeting is the envy of every other subfield, and our students do extremely well on the job market. In every way we are one of the most, if not *the* most, accomplished groups in political science. These are achievements that we should both be justifiably proud of and appreciative for the excellent leadership that we have enjoyed over the last quarter century.

Naturally, some challenges remain. We are experiencing growing pains that are unprecedented in the Society's history. Clearly this is a good problem to have but it leads to difficult decisions with regard to our Summer Meeting as well as issues of representation. Ours is a technical discipline, but one that feels a natural inclusiveness for all that choose to work hard to obtain the requisite skills and abilities. In fact, I would argue that we are the most egalitarian subfield of an often prejudiced field. So we should continue to find a way to make the Summer Meeting, conference panels, and all other outlets as open as possible to scholars who contribute at an expected high level. Relatedly, securing venues for the Summer Meeting continues to be a challenge. However, we are lucky to have the next three lined up (University of Iowa 2010, Princeton University 2011, and the University of North Carolina & Duke 2012). In the future we will work to make it as easy as possible for interested members to convince their administration that it is both good public relations and good economic sense to host. Growth is good, and we are second in APSA section numbers behind only the omnibus juggernaut of comparative politics (see Figure 1).

We also enjoy a privilege rare to other political scientists. Increasingly, other academic fields in the social sciences and elsewhere are turning to us for methodological guidances. Impressively, I am aware of a number of our colleagues that participate in medical studies in a leadership capacity. It is clear me, and to others, that Political Methodology is on the forefront of quantitative research and understanding amongst the social sciences (including being ahead of that discipline that starts with an "e"). So we are not only the most sophisticated quantitative researchers in the social sciences, we rank amongst the most technically accomplished in any discipline.

Some recent achievements by past Presidents deserve mention. Phil Schrodt has initiated a wonderful project designed to make curricular materials widely available and free for our students. Jan Box-Steffensmeier almost singlehandedly reaffirmed our commitment to the inclusion of under-represented groups, women in particular. Both of them were critical in the last renewal of our National Science Foundation grant to support graduate student attendance at the Summer Meeting.

Figure 1: Comparison of Section Membership



This naturally leads to the question of what I would like to initiate during my brief period as President. As a precursor, let me say that my first priority is on getting the National Science Foundation grant renewed so that we can continue to do the important work that this support has allowed. In addition, the section officers and the publications committee will ensure a smooth transition to a new editorial team and a new or renewed publisher for *Political Analysis*. We already have a new team in place to run *The Political Methodologist*: Wendy Tam Cho, Jake Bowers, Jude Hayes, and Brian Gaines, all at the University of Illinois, Urbana-Champaign. Having noted these issues, I feel that our best opportunity for development lies in internationalization. Right now 22% of our section members are from 36 countries outside of the United States (223/1027 of dues paying members; 1880 subscribe to the polmeth list but the complexities of email addresses make it impossible to breakdown this other number). As I am sure you know, political science is an incredibly American-centered discipline. I'll avoid the difficult and normative discussion of the causes and cures for this. Importantly, though, we are the most obvious subfield for international outreach. Our core materials and ideas supercede linguistic differences in the same way the fields of physics, chemistry, and engineering do. Therefore there is no more logical group to reach out across the globe from the discipline of political science.

Saying that we should internationalize and doing so are of course radically different. More concretely, I suggest that we do three things. First, we should endeavor to financially support international attendees to our Summer Meeting. This clearly has a wide range of positive effects. Secondly, I propose that we develop a Spring Meeting outside of the United States. A number of institutions have already expressed interest. This should not be seen as replacing or eclipsing our flagship Summer Meeting, but rather as an evangelical opportunity. I recognize that the financial burdens of traveling to Europe, Asia, or elsewhere, on top of current conference commitments may be prohibitive for some, but the advantages for spreading our gospel make this well-worth the effort. To belabor the analogy, the power of our message is strong and we should make it readily available as widely as possible. Finally, international submissions to *Political Analysis* are critical in keeping up our commitment to like-minded scholars all over the world.

So at the end of the day (or the end of the article) we are in a very healthy position, admired by our colleagues outside of the subfield, financially secure within the subfield, and able to take on new challenges. I look forward to working with you over the next two years. Please feel free to contact me with any suggestions or comments that you might have.

Jeff Gill

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. 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 LaTeX on a Mac OS X running iTeXMac.