LAB NOTES

Vol. III, No. 1 (Autumn 2003)

The Official Quarterly Newsletter of the Political Research Laboratory

Director's Welcome

by Herb Weisberg

Autumn is always an exciting quarter, with people returning to school and new people joining us – and this autumn is especially exciting for the Political Research Laboratory (PRL or Lab) with a new staff person joining the operation, a new program in the Lab, and new Associate and Assistant Lab Directors, along with new graduate and undergraduate students working in the lab.

The new staff person is Bill Miller, who will be joining James Norman on the computer side of the Lab. Bill is an experienced OSU hand, previously being one of the senior people at the OSU Computing Center and at the OARNet supercomputer operation. As he learns more about our operation, he will be sharing a full set of computer responsibilities with James. In particular, he will be doing morning computer hours, helping the front office on programming applications, working on the department web page, and being in charge of inventory. Be sure to come by, meet Bill, and welcome him to the department.

The new program in the Lab is PRISM: the PRogram In Statistical Methodology, led by Jan Box-Steffensmeier as the new Associate Director of the Lab. The Lab has long given statistical advice to faculty and grad students, but PRISM provides a welcome new opportunity to give further emphasis to this important role. Kevin Sweeney is moving over to be the Senior PRISM fellow, and Brandon Bartels decided to be Junior PRISM fellow so that he wouldn't have to move offices – welcome aboard, Brandon.

The new Assistant Director of the Lab is Greg Miller, who you all should know as a senior graduate student in International Relations. He will be responsible for the day-to-day operations of the Lab, making sure that everything goes smoothly.

The new graduate assistant in the Lab is Zach Mears, who will be joining Brent Strathman to work on computer trouble-shooting, the department web page, the human subjects pool, and so on. The new undergrad is Michael Kovalchuk, a talented sophomore in Computer and Information Sciences. Brent, Zach, and Michael are the first level of contact for when you have computer problems – at least one of them will be in the lab between 8:30 a.m. and 6:00 p.m. Monday through Friday.

When you drop by the Lab, you will find a considerable shifting of offices – I am back in the east-most end office, Jan is in the west-most end office, Bill Miller is in the old conference room,

the conference room is in the old high-tech room, and the high-tech computer (a machine fully loaded for statistical applications) and the scanner computer are in one of the cubbies (2049N). And we are working to make the Archive Room more useful by arranging some codebooks more thematically, such as separate shelves for American National Election Studies, for Eurobarometers, and so on.

We are working hard to improve the Lab's infrastructure, including a new switch that should increase the web security of the department's Windows 95 and 98 computers and a new round of computers for the Experimental Lab.

Meanwhile, I remain surprised to be back in the Lab operation. I became Director the first time exactly ten years ago, when the department moved back from its exile in Neil Hall to the then freshly renovated Derby Hall. A lot has changed for the Lab since then, with the departure of the survey unit, the arrival of the experimental lab and the new PRISM program, the retirement of Jim Ludwig, and now the addition of Bill Miller. What remains unchanged is our desire to be maximally helpful to the department's faculty, staff, and graduate students in computer and research support and in facilitating high-tech teaching. Be sure to let us know if you see ways that we can be more useful to supporting the department's research and teaching missions.

PRL Services or: How I Learned to Stop Worrying and Love the Lab

For many of you, the PRL is simply the room with the computers. While maintaining the common computer and printer area is an important component of the Lab, it is not the PRL's only function. The following is an overview of just some of the services we provide to the department's students, staff, and faculty:

Computers, Networks, and Email

The Lab sets up departmental email accounts. The PRL also maintains and upgrades the department's computers and the network. If you have any trouble with your email account or with your office equipment, be sure to contact the Lab. The PRL also houses the stats machine, a high-tech

computer, for running analyses of very large datasets. This machine, along with the scanner, can be found in 2049N.

Department Website

We are in the early stages of revamping the department's website. Graduate students and faculty members have their own web pages, but the challenge has been to allow individuals to update their personal pages while preserving the professional look of the department's website. The next year will see a number of changes to the website, in terms of its structure, the information available, and the degree to which individuals have access to their departmental pages. If you have any suggestions for the website, please email them to the PRL at prl@polisci.sbs.ohio-state.edu.

Experiments and Subject Pool

Last year was a successful year for the subject pool and the experimental lab, as hundreds of subjects participated in over a dozen experimental runs. If you are an autumn instructor looking for ways to give students extra credit, the subject pool will be running throughout the term. Contact Zach Mears at mears.16@osu.edu or 292-0511 for more information or to sign up.

In the future, the Lab hopes to integrate information on past and current experiments to make the subject pool process easier for participants and instructors. Furthermore, we wish to create a clearinghouse of experimental data sets and papers that made use of the experimental lab. Both will be accomplished with the creation of a new web page devoted to experimental methodology. Finally, we hope to make the experimental lab more widely available through brown bags on methodological and practical issues in experimentation.

Equipment Loans

The Lab has four laptops and one data projector to be loaned out. We do our best to accommodate all requests, however demand is high, particularly for the lone projector. Therefore, we ask that you give at least one week's notice if you are going to need any

equipment. Contact Greg Miller at miller.2452@osu.edu or 292-1061 with requests.

Brownbag Meetings

The Lab frequently sponsors informational sessions to assist in the department's development. Over the next year, the Lab staff looks forward to providing brownbag meetings on the use of Microsoft PowerPoint in the classroom (a beginning level session in the Fall as well as a more advanced session in the Winter or Spring) and how to develop your departmental webpage. If there are additional topics you would like to see addressed, please contact Greg Miller at miller.2452@osu.edu or at 292-1061.

Data Retrieval

Previously, data from the Inter-university Consortium for Political and Social Research (ICPSR) was obtained through the PRL. However, with the advent of modern technology, students and faculty may download data from www.icpsr.org using their OSU username and password. If you have any problems obtaining the data, the PRL is still available to help. Questions should be address to Kevin Sweeney, who has taken on data retrieval responsibilities as part of the PRISM program.

PRISM: PRogram In Statistics and Methodology

The creation of PRISM, which is tied to the PRL, launches a renewed focus and commitment to methods training in the department. We hope to create a close community of faculty and graduate students who are interested in methods, and to move beyond the department by fostering ties with Sociology and Biostatistics. A PRISM speaker series will be launched this fall, to include Paul Kellstedt, Texas A&M University, on November 13 & 14 (http://www-polisci.tamu.edu/faculty/byname/?id=67), as well as Narayan Sastry, RAND Institute (http://www.rand.org/labor/staff/sastry/profile.html). PRISM also includes two 12-month positions for a senior and a junior methods fellow.

Kevin Sweeney currently holds the Senior Methods Fellowship and his duties include:
1. Providing technical assistance to advanced methods classes and visiting each class at least once, presenting a lecture on a specific methodology or program at the discretion of the instructor and the PRISM Director.
2. Presenting his own research during the year for the PRISM Paper Series, and is to be a regular contributor to the departmental Methods Luncheons. To that end, the Fellow should get to know those in the Department who are interested in Political Methodology, and be available as a resource to them.
3. Holding regular consulting hours where advanced graduate students and faculty may come to ask questions, discuss various methods and data analysis problems. Of course there are limits due to time constraints (this is a fellowship afterall), but Kevin is an extremely valuable consulting resource so come to him with your hard questions.
4. Must be skilled at data acquisition and manipulation so that he or she may be a resource to those in the department that have such problems.
5. Following consultation with the PRISM director, the Fellow will be mentored by and strongly encouraged to collaborate with OSU faculty member(s) on a research paper. Kevin, Brian Pollins, and Omar Khesk are collaborating on a project examining cooperation within the OPEC oil cartel.

Brandon Bartels hold the Junior Methods Fellowship and his duties include:
1. Providing technical assistance to the students and instructors in methodology courses.
2. Organizing the PRISM luncheons, be available to interested participants as a resource to them, and help cultivate methods interests for the first year students.
3. Holding regular office hours where graduate students may come to ask methods questions. It is strongly recommended that those with questions come prepared with a "game plan"/specific questions, and emailing either of the methods fellows in advance is useful for all involved.

4. Following consultation with the PRISM director, the Fellow will be mentored by and strongly encouraged to collaborate with OSU faculty member(s) on a research paper. Brandon and Jan Box-Steffensmeier have two projects underway. The first uses a split-population duration model in the area of campaign finance. The second employs a new panel estimator for discrete data that allows for both dynamic effects and individual heterogeneity. The paper uses individual level party id data.

We expect there to be a close mentoring relationship between the fellows and faculty. Indeed, we hope that by creating PRISM, there will be a closer mentoring relationship in the area of methods between all graduate students and faculty with methods interests. We see the Methods Fellows as a collective good via their consulting roles for the department. PRISM will be continuing the departmental methods luncheons as a key to fostering community building – please join us. We plan to have presentations on methods topics and software, speakers from the department and outside the department, discussions of articles, and problem solving sessions. Please let Brandon know if you would like to present, have an article you would like to discuss, or have other suggestions for the luncheons or more generally, for PRISM.

Computer Viruses

Just a quick reminder about computer viruses, worms, and other nasty stuff…

Be cautious when opening any file that is attached to an email. A general rule of thumb is to not open any attachments from someone you do not know. However, it is possible for someone to emulate a sender's email address so that it appears as if it comes from a friendly source. One common virus comes as an attachment to an email, claiming to be a security patch from security@microsoft.com.

BRINGING CIVILIZATION TO ITS KNEES…
Goths / Vandals / Huns / Geeks

The moral is basically to be vigilant; if you are in doubt about the validity of an email attachment, ask someone in the Lab.

Advances in Duration Modeling: The Split Population Duration Model
By Brandon Bartels

In this Statistics Corner, I will review an exciting and important development in duration modeling—the split population duration model—which is applicable to many important questions in political science. This model accounts for a particular type of heterogeneity across observations, namely, it relaxes the assumption that all subjects will eventually experience the event of interest. Before discussing this model, I present a brief overview of duration models for political science.

Duration Models in Political Science

Duration models allow researchers to tackle important theoretical questions concerning the timing of events or the survival of particular states of being. The study of *when* events occur is often at the center of political inquiry, and the importance of timing in politics takes us back to Fenno's contention: "If we are to explain outcomes, who decides *when* may be as important as who decides *what*" (1986, 9). Indeed, in the past decade, scholars have used duration models to address interesting questions of timing and survival, such as the dissolution of cabinets in parliamentary governments (King et al. 1990;

Warwick 1992; Diermeier and Stevenson 1999), the duration of wars (Bennett and Stam 1996), the survival of political regimes (Bueno de Mesquita and Siverson 1995), the timing of state policy adoption (Berry and Berry 1990; Volden 2003), challenger entry in congressional elections (Box-Steffensmeier 1996), the timing of position-taking in Congress (Box-Steffensmeier et al. 1997; Caldeira and Zorn 2003), delay in Senate confirmations of presidential nominees (McCarty and Razaghian 1999; Binder and Maltzman 2002; Martinek et al. 2002; Shipan and Shannon 2003), and the survival of precedent in the U.S. federal courts (Spriggs and Hansford 2001; Benesh and Reddick 2002).

While interest in questions of timing unifies scholars using duration models, some differences remain within the community, primarily those centering on how to treat the *hazard rate,* or a subject's risk of experiencing the event of interest, given that it has not yet done so.[1] Parametric models specify explicitly the distribution of the hazard; examples of parametric distributions include the Weibull, exponential, and Gompertz, as well as accelerated failure time (AFT) formulations such as the log-logistic and log-normal. Parametric models, then, explicitly account for *duration dependence,* or the extent to which the risk of experiencing the event increases or decreases as a function of time. On the other hand, the Cox model—the semi-parametric approach—leaves the baseline hazard unparameterized, which means that the hazard is not constrained to possess a specific distributional form. Thus, the Cox model estimates the effects of independent variables on the timing of the event of interest, and leaves duration dependence unspecified, essentially treating it as a nuisance factor. The choice of parametric versus semi-parametric models is typically left to the researcher's confidence in the theory underlying the data-generating process. If one possesses strong theory as to the distributional form of the hazard, then using a parametric model may be appropriate. However, if one is uncertain about the shape of the underlying hazard rate for the process under study—indeed, many scholars remain skeptical that social science theory is precise enough to justify a

parametric model—then the Cox model provides added flexibility over parametric models. Readers with further interest in the details surrounding parametric versus semi-parametric models should consult more general duration modeling sources (e.g., Box-Steffensmeier and Jones 1997, 2003; Blossfeld and Rohwer 1995; Hosmer and Lemeshow 1998).

The Split Population Duration Model

Split population duration (SPD) models account for a specific type of heterogeneity, i.e., the possibility that some cases will never experience the event of interest while some will. One of the assumptions of standard duration models is that every observation in the data *will eventually experience the event of interest*, which is sometimes an unreasonable assumption in violation of a particular theory or understanding of the process under examination. Developed in biostatistics about 50 years ago and popularized in economics and sociology by Schmidt and Witte (1984, 1989), SPD models relax this assumption by essentially "splitting" the observations under analysis into two subpopulations, one that will eventually experience the event of interest and one that will never experience the event. In their study of criminal recidivism, Schmidt and Witte (1984, 1989) were interested in the factors explaining the timing of criminals returning to prison after being released; but we know that not all criminals return to prison after release, a factor incapable of being accounted for by a standard duration model. To solve this problem, Schmidt and Witte specified a model that generates two sets of simultaneously estimated coefficients: one for the likelihood of the event ever occurring and the other for the timing of the event, conditional on the event ever occurring. The mathematical derivation of the model, drawing on Schmidt and Witte (1989) and Box-Steffensmeier et al. (2003), is located at the web archive (see below for website information).

It is worth emphasizing a few important points about the SPD model. First, two sets of coefficients are estimated in SPD models: (1) coefficients for the effects of covariates on the *incidence* of the event occurring, and (2) coefficients for the effects of covariates on the

---

[1] Functionally, the dependent variable in duration models is the time until the event of interest occurs. Conceptually, though, the dependent variable is actually the unobserved hazard rate.

timing of the event, *conditional on the probability the event occurring*. It is also important to underscore that the censoring indicator (i.e., whether or not we observe the event occur within the analysis time) serves as the dependent variable in the incidence portion of the model. Second, a very powerful feature of these models is that different covariates can be included to explain *whether* and *when* the event occurred. For example, an independent variable may have a positive effect on *whether* the event occurred and a negative effect on *when* it occurred. This makes the split population model much more flexible than other duration models where the effect of timing and incidence are combined. Third, the SPD model estimates a "split parameter," δ, which is the estimated mean probability of cases experiencing the event of interest. This statistic allows the analyst to test whether relaxing the assumption that every observation will experience the event of interest is necessary. If it is not, i.e., if $\delta = 1$, the SPD model collapses into a typical duration model. So there is little cost to estimating the split population model, and as such, scholars should almost always use the SPD model when they have reason to believe that not all observations will experience the event of interest. The estimated split also serves as a sort of goodness-of-fit statistic in that it can be compared to the proportion of cases that actually experienced the event of interest. Fourth, SPD models are currently only estimable using parametric approaches. For those who advocate the semi-parametric approach to duration modeling, this is certainly the downside of the SPD model. However, work continues to be done to estimate a Cox-type SPD model, although problems of model identification have hampered these attempts so far (see, e.g., Sy and Taylor 2000; Kuk and Chen 1992). Fifth, regarding software capabilities, LIMDEP is currently the only software that has a canned routine for estimating SPD models. In addition, Forster and Jones (2001) have written programs for estimating these models in Stata (see the web archive for an example).

Finally, an exciting and relatively new development includes the capability of SPD models to generate estimates for time-varying covariates (TVCs). For many scholars, one of the disappointments of the Schmidt and Witte model was that it could only estimate the effects of time-invariant covariates. However, Forster and Jones (2001), in an effort to study the effect of increases and decreases in taxes on the timing of people starting smoking, have developed an SPD model to accommodate TVCs, providing even more flexibility to the model.

Applications

Surprisingly, very few applications of the SPD model exist in political science. In fact, I am aware of only one published paper, by Clark and Regan (2003), using an SPD model. In that paper, the authors analyze the timing of interstate conflict, and they account for the notion that not all dyads have the "opportunities" to engage in war. Two other unpublished papers use SPD models to study the incidence and timing of events. First, Box-Steffensmeier et al. (2003), recognizing that early money in campaigns influences the success of campaigns, examine the factors that affect the timing of PAC contributions to incumbent U.S. House members. Using datasets of incumbent-PAC dyads for both labor and corporate PACs, the authors argue that the use of a standard duration model would be unreasonable, since it would assume that both labor and corporate PACs would eventually give to *every* incumbent. From the literature on PAC contributions, we know that there are certain Representatives who would never receive contributions from labor or corporate PACs. For instance, we should not expect labor unions to contribute to Rep. Peter Hoestra, who has investigated the Teamsters, or Cass Ballenger, who has tried to reform OSHA. Thus, Box-Steffensmeier et al. use the SPD model to examine the factors that explain both the incidence and timing of PAC contributions.

Another application of the SPD model is by Hettinger and Zorn (2001), who analyze the timing of congressional overrides of Supreme Court decisions. A standard duration model would assume that all Supreme Court cases will eventually be overturned by Congress, certainly an unrealistic and unreasonable assumption that defies much of what we know about the survival of the Court's decisions. Thus, Hettinger and Zorn specify an SPD model estimating the effects

of independent variables on both the incidence and timing of congressional overrides. In interpreting the timing coefficients, the effects are conditional on the probability of the case ever being overturned. Also, the estimated split parameter of .09 means that the model predicts that 9% of the cases will actually be overturned, certainly justifying the use of the SPD model. The actual percentage of cases overturned in the sample is 7%, indicating that the model does a reasonably good job in "splitting" the population. Certainly, a standard duration model, which would imply a splitting parameter of 1.0, would produce incorrect estimates of the effects of variables on the timing of overrides (see Hettinger and Zorn for a comparison of the SPD model with a standard model).

Conclusion

Given its capability to incorporate important information about the likelihood of different cases experiencing the event of interest, the SPD model certainly represents an exciting advance in duration modeling. With the recent development of SPD models to include TVCs, the constraints of using the SPD model have been reduced even more. However, those who advocate the use of semi-parametric duration models will undoubtedly be hesitant to use the SPD model, which currently can only accommodate the parametric approach. But in general, scholars who have reason to believe that not all observations will experience the event of interest should almost always estimate an SPD model instead of a standard duration model.

Web Archive

To view the mathematical derivation of the SPD model, an SPD program written for Stata, and the references cited in this Statistics Corner, go to http://psweb.sbs.ohio-state.edu/grads/bartels/statscorner.htm

FAQs for First Years (and for those of us who haven't been paying attention)

Where do I go when I have a computer-related problem?

At least one member of the PRL staff is "on-call" from 8:30 a.m. until 6:00 p.m. Monday through Friday; the on-call schedule is posted in the computer lab. Otherwise, you can contact us using the information on the next page.

Why do I have two different email accounts?
Everybody has a department email account and a university email account. You may have your mail forwarded from one of these accounts to the other, or have both forwarded to your personal email account (e.g., hotmail, yahoo, aol, etc.). Contact the PRL to have your department email forwarded. The Office of Information Technology (OIT) is responsible for university accounts; forward university email using this online form: http://8help.ohio-state.edu/mail_forwarding.html.

Is there an easy way to transfer files from one computer to another, without needing a disk?
Yes. The I: drive is a public drive. Anything you save to that drive can be accessed from any computer (but can also be accessed by anyone else). In addition, everyone has an account on the K: drive, which is also accessible from every computer, but only by the account holder (you must log into the network to access the K: drive)

Why does my graduate student office computer only have a limited amount of software?
Because the department must purchase licenses for any software we use, we are limited in what we can provide to everyone. Generally, we make sure each office has the essentials, and any other software that is frequently used will at least be loaded on one of the Lab computers. If you do have any special software requests, please notify the PRL. Some software can be downloaded for free from the university. See the OIT website at: http://osusls.osu.edu/

Where are the fastest computers that graduate student can use?
Besides the stats machine in 2049N, the fastest computers are those in the experimental lab (2049A). Graduate students are welcome to use these computers when no experiments are

underway. Under university rules, we are unable to purchase new computers for graduate student offices, but we hope to upgrade the computers in other lab locations this year.

Why aren't all of the walls in my office smooth?
Because those walls have asbestos in them – which means that you should not put nails or screws through walls that have unsmooth surfaces. Reassuring, isn't it?

Why is there never paper in the computer lab's printer?

To save money, we ask that students provide their own paper for printing in the PRL. In addition, we ask that large print jobs (such as JSTOR articles) be kept to a minimum, especially when others are waiting to use the printer.

Why did I spend all this time reading LabNotes?
Just think of all the crucial knowledge you would have failed to obtain, all of which will be relevant during your stay in Derby Hall. Plus, it was a nice diversion from King, Keohane, and Verba wasn't it?

PRL Staff Information

Title	Name	Office	Phone	Email
Director	Herb Weisberg	2049D	292-6572	weisberg.1@osu.edu
Associate Director	Jan Box-Steffensmeier	2049S	292-9642	jboxstef+@osu.edu
Assistant Director	Greg Miller	2049R	292-1061	miller.2452@osu.edu
Systems Manager	James Norman	2049K	688-3732	norman.67@osu.edu
Systems Manager	Bill Miller	2049E	292-1814	bmiller@polisci.sbs.ohio-state.edu
Graduate Assistant	Brent Strathman	2049F	292-0511	strathman.5@osu.edu
Graduate Assistant	Zach Mears	2049F	292-0511	mears.16@osu.edu
Undergraduate Assistant	Michael Kovalchuk	2049H	292-0511	kovalchuk.2@osu.edu
PRISM Senior Fellow	Kevin Sweeney	2049Q	292-9661	sweeney.101@osu.edu
PRISM Junior Fellow	Brandon Bartels	2049Q	292-9661	bartels.20@osu.edu

To contact all Lab Staff simultaneously, email us at prl@polisci.sbs.ohio-state.edu

Previous issues of Lab Notes and other valuable information can be found at the Lab's website:
http://psweb.sbs.ohio-state.edu/prl/index.htm