

Qualitative & Multi-Method Research

Newsletter of the
American Political Science Association
Organized Section for Qualitative and Multi-Method Research

Contents

Institute for Qualitative and Multi-Method Research Tenth Anniversary

“Boot Camp”: Ten Years of Qualitative and Multi-Method Research
 Colin Elman2

Retrospectives from Attendees
 Pam Bromley, Stephanie McNulty, Sarita D. Jackson, Christopher Darnton, Nivien Saleh, Halit Mustafa Tagma, Maria Koinova, Lily Hsueh, Juan A. Bogliaccini, Catherine Herr-old, Steven Samford, Kate Svyatets, Ajay Verghese, Kerry Crawford, Anastasia Shesterinina, and Patrick Mello5

Symposium: Methodological Issues in Medium-N Research

Between a Rock and a Hard Place: The Uneasy Position of Medium-N Research in Comparative Political Science
 Patrick Emmenegger14

Case Selection in Medium-N Comparative Welfare State Research
 Bernhard Ebbinghaus15

One-Size-Fits-All? Measurement Issues in Medium-N Comparative Welfare State Analysis
 Jon Kvist20

Step Away from the Pool
 Lane Kenworthy26

Fast-Moving and Slowing-Moving Factors: Taking Apart Time-Series Cross-Sectional Analysis
 Gary Goertz28

Political Science Seen Through Historians’ Looking Glass: The Case of Electoral System Choice
 Patrick Emmenegger and Klaus Petersen33

Regime-Based Analysis in Comparative Welfare State Research: Old and New Approaches
 Michael Shalev37

Announcements

Letter to the Editor
 Kevin Bruyneel, Lee Ann Fujii, Patrick Thaddeus Jackson, Xymena Kurowska, Cecelia Lynch, Ido Oren, Timothy Pachirat, Edward Schatz, Ronald Schmidt, Sr., Peregrine Schwartz-Shea, Joe Soss, Dorian Warren, Cai Wilkinson, Dvora Yanow, and Cyrus Ernesto Zirakzadeh43

Section Awards45

Letter from the Section President

Gary Goertz
 University of Arizona
 ggoertz@u.arizona.edu

As incoming president of the section I would like to take this opportunity to thank my predecessor Colin Elman for his service to the section. He was instrumental in setting up the David Collier Mid-Career Achievement Award as well as a new *APSR* qualitative paper submission award. As most of you know the section would not be what it is without his large investment of time and energy. So on behalf of the section: Thanks, Colin!

As many of you know the *APSR* is changing editors. I was appointed to the selection committee, presumably to represent the section. I do not believe I am violating confidence to say that a key part of our charge from the APSA Council was to evaluate candidates based on their ideas and commitment to increasing diversity in the *APSR*. Diversity has many dimensions, but one is certainly methodological. One of the key issues was, “What is *APSR*-level qualitative research?” Since the *APSR* has published virtually no qualitative articles, this is an unknown standard. I encourage members of the section to help define that standard by submitting their work to the *APSR*. One question I asked many candidates was whether work good enough for the best subfield journals of international relations and comparative politics would be worthy of very serious consideration at *APSR*. The answer from most candidates was yes. So if you think your work is good enough for major subfield journals I encourage you to submit it to the *APSR*. If you do decide to submit please also email me your submission for consideration for the section’s *APSR* paper submission award. Details of the award can be found in the call for nominations, which is reprinted on p. 47 of this newsletter.

APSA-QMMR Section Officers

President: Gary Goertz, University of Arizona
 President-Elect: Lisa Wedeen, University of Chicago
 Vice President: Jeff Checkel, Simon Fraser University
 Secretary-Treasurer: Renee de Nevers, Syracuse University
 QMMR Editor: Robert K. Adcock, George Washington Univ.
 Division Chair: David Waldner, University of Virginia
 Executive Committee: Elisabeth Wood, Yale University
 Stephen Hanson, University of Washington
 Giovanni Capoccia, University of Oxford
 Benjamin Smith, University of Florida

Institute for Qualitative and Multi-Method Research (IQMR) Tenth Anniversary

“Boot Camp”: Ten Years of Qualitative and Multi-Method Research

Colin Elman

Syracuse University
celman@maxwell.syr.edu

The essays in this symposium reflect on how the Institute for Qualitative and Multi-Method Research (previously the Institute for Qualitative Research Methods) influenced the authors' respective trajectories. This short introduction describes the genesis of the institute, and discusses some of the ways that it has changed over the ten years it has been held.¹

The original idea for the institute arose from the mismatch between the wide use of qualitative methods in political science research, and the scarcity in U.S. political science departments of specialized graduate classes teaching these methods (see, e.g., Bennett, Barth, and Rutherford 2003). Even where qualitative methods were taught, they were often treated as just one component of a broader scope and methods course, rather than given sustained coverage. One early manifestation of the push for more and better training was a suggestion by Alexander George to Andy Bennett to develop a traveling workshop which would visit different departments and provide a short introduction to qualitative research methods. Separately, Stephen Walker, then Chair of the Department of Political Science at Arizona State University, encouraged me to organize an institute to be hosted in Tempe.² Andy was amenable to our joining forces, and we began planning in 2000 for the first institute to be held in January 2002. David Collier soon joined the enterprise, and the three of us have been organizing the institute since then with a core group of repeating instructors and a wider group of faculty whom we have periodically invited to the institute to broaden its curriculum.

Despite our involvement in the genesis of the institute, David, Andy, and I strongly resist the characterization of the institute as being the product of our efforts. Our initial moves to organize the institute pale in comparison with the sustained commitment from its instructors and participants. With respect to the faculty, all the instructors play a role in determining what gets taught at the institute. They each have considerable autonomy in how they organize and run their sessions. The instructors are experts in (and often the authors of) the material being taught. In addition, they are teaching what they love, and their passion has been one of the most compelling virtues of the institute. One manifestation, latent to anyone but the host institution, is that the instructors have been mostly indifferent to whether they were paid for their efforts. (For the record, I can say that they were paid, although not as generously as we would have liked or as well as they deserved.) The central point is that out of the more than 250 teaching slots at the first ten institutes, I can count on the fingers of one hand the times an instructor asked how much the teaching honorarium was

before accepting an invitation to help lead the institute. They just wanted to teach the material.

If the instructors' expertise and passion provide part of the institute equation, the energy and enthusiasm of the participants were likewise essential. The first email suggesting the institute had the subject line “Boot Camp,” and it was an accurate prediction of what ensued. The formal timetable of the institute runs from 8:30 am to 5:30 pm for five days; there is then a two-day break, and then it repeats for five more days. This does not take into account the heavy reading load to prepare for the daily sessions, or the careful parsing of fellow participants' research designs, which receive critiques at the institute. The nonstop pressure makes it all the more remarkable that it is commonplace to the point of cliché that on any evening of the institute there will be groups of participants in the hotel bar or lobby discussing a research design that was covered earlier that day, arguing the merits of a methodological disagreement between two instructors, or planning a conference panel or joint journal submission.

The commitment from instructors and participants has been remarkable, and thankfully has not changed despite the growth of the institute. The first gathering took place at Arizona State University in 2002 and, as noted in Table 1, involved 45 students and 18 faculty. Over the years, the institutes have grown larger, and in 2011 we welcomed 159 participants and 30 faculty. In the ten years the institute has been in operation, it has hosted 1,086 students, taught by 267 faculty.³

There are two ways to attend the institute: by nomination from a member of the Consortium for Qualitative Research Methods, or by successfully competing through the open application pool for a fellowship slot.⁴ Membership in the Consortium is at the unit level (i.e., department, research center), and so there have often been universities with multiple subscriptions. For example, at various times more than one unit has signed up from Princeton University, Yale University, Harvard University, Georgetown University, Indiana University, and the University of California, Berkeley. There are three levels of subscription, each with a different number of nominations to the institute: affiliate carries one slot, sponsor two, and premium sponsor three. Nomination includes tuition, lodging, and board for the duration of the institute. The number of member units has grown steadily over time. The Consortium began with 20 members, and in 2010–2011 it had 72 subscribers (see Appendix 1 below for a list of universities with one or more member units between 2009 and 2011). Although the institute is mainly funded by member subscriptions, it has also received substantial support from other sources. This has included initial seed funding from Arizona State University, current support from the Moynihan Institute of Global Affairs at Syracuse University; and six grants from the National Science Foundation totaling \$345,020.⁵

The second way to attend the institute, through an open pool application, involves competing for a fellowship that covers board, lodging, and tuition for the duration of the institute.

Table 1: Institute Summary Data, 2002–2011

Institute Year	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	TOTALS
# of institutional subscriptions to CQRM	20	26	28	35	48	52	62	66	64	72	473
# of instructors	18	22	27	25	28	30	26	32	29	30	267 ^a
# of students attending Institute	45	63	70	88	105	127	129	153	147	159	1,086
# (%) of female students	21 (47%)	28 (44%)	39 (56%)	51 (58%)	54 (51%)	62 (49%)	71 (55%)	85 (55%)	69 (47%)	72 (45%)	552 (50%)
# of open-pool applications	198	145	69	103	140	135	134	198	212	241	1,575
# of open-pool slots	13	9	13	9	6	9	9	9	12	12	101
Percentage of open-pool applicants accepted	7%	6%	19%	9%	4%	7%	7%	5%	6%	5%	6%
Percentage of Institute attendees from open pool	29%	14%	19%	10%	6%	7%	7%	6%	8%	7%	9%

Note: (a) Total number of instructor slots; includes some faculty who repeated.

The open-pool competition is for those who either are enrolled in member institutions but are not nominated by them, or who are from non-member institutions. The pool is more or less self-generating. Beyond a couple of email announcements, we do no advertising of the institute (we produce no brochures nor do any mailings); instead, much of the “promotion” occurs through word-of-mouth by previous attendees. Demand for the open-pool slots has grown significantly over the last few years, to 241 applications for the 2011 Institute, roughly 20 applications for each available slot. The criteria for selecting students from the open pool include: (1) overall record, as indicated by the student’s c.v.; (2) quality of personal statement on why they wish to attend the Institute; (3) stage in academic career (we slightly favor junior faculty over graduate students and prefer to accept advanced graduate students who are working on their Ph.D. dissertations); and (4) ethnic, gender, and geographic diversity. Although we are pleased that the open pool has become so popular, we deeply regret having to turn away so many well-qualified applicants. We hope to be able to do better in 2012, since we are increasing the number of open-pool slots from 12 to 18.

The institute has attracted a diverse group of scholars from across the sub-fields of political science and its cognate disciplines. Women have made up 50% of all attendees. These are extraordinary participation rates compared with other methodological programs. For example, less than 20% of those who attended the Society for Political Methodology’s annual summer conference between 1984 and 2010 (466 of 2,365 attendees) were female.

Over the years, the core focus of the institute—intensive instruction in qualitative/multi-methods approaches—has remained constant. The institute is designed to enable graduate students and junior faculty to create and critique methodologically sophisticated qualitative research. Institute classes

explore the uses, techniques, strengths, and limitations of qualitative methods. Topics have included research design, concept formation, methods of structured and focused comparisons of cases, typological theory, case selection, process tracing, comparative historical analysis, congruence testing, path dependency, interpretivism, counterfactual analysis, field research techniques (including interviewing and archival research), necessary and sufficient causation, fuzzy-set methods, and philosophy of science issues relevant to qualitative research.

However, while the core focus has remained steady, the expansion in the size of the institute has motivated changes in the institute’s structure. In particular, the growth has been managed by moving to a modular format. This has also allowed us to respond to the dramatic enlargement in the core qualitative and multi-method canon, which has grown both in terms of volume and diversity of content. We continue to offer sessions that all participants are required to attend on foundational issues of epistemology and causation, and on obtaining funding for and publishing qualitative and multi-method research.⁶ But eight of the ten days of the institute now offer choices of either two or three concurrent modules, each comprised of three sessions. The modular format allows us to offer instruction in methods drawn from across the qualitative and mixed-methods canon while keeping class sizes reasonably small. It also gives students control over their options, allowing them to select techniques that will most benefit their work.

Any history of the institute, however abridged, would be woefully incomplete without mention of the research design discussion groups. Participants in the institute are strongly encouraged to prepare a written research design in advance, and sessions focused on the research designs are one of the most valuable parts of the program. On most days, small

groups come together for the presentation, discussion and (constructive) critiquing of the research designs, and to receive feedback from fellow participants and at least one faculty member. These sessions help to reinforce many of the methodological ideas explored in the lectures and break-out sessions, showing how these ideas can be applied in ongoing research. Almost all of the attendees take advantage of the opportunity to present their own research designs in small-group discussion sessions. Institute organizers seek (not always successfully) to structure these sessions such that the three research designs presented in each are on related themes. The sessions are formatted as seminars with one faculty member serving as chair, and roughly fifteen participants in attendance. Students consistently offer each other intensive, thoughtful, and constructive criticism and feedback. Participants are expected to attend seven of these 120-minute sessions during the two-week institute, and thus engage in the discussion of 21 designs (their own plus 20 others).

Apart from the change to a modular format mentioned above, the other major structural change at the institute has been the addition of an authors' workshop. The institute had always provided a forum for methodologists—including institute faculty, as well as some student/participants with advanced methodological skills—to discuss their own writing and publications. Early versions of books and articles developing, disseminating, and/or using qualitative methods have been presented and discussed at the institute, giving the authors an unparalleled opportunity to “road-test” forthcoming work on a multi-subfield audience, and to receive feedback and criticism from students and instructors alike. In addition, conversations at the institute resulted in collaborative writing projects among faculty and catalyzed new methodological research by faculty and students alike. With some timely encouragement from Brian Humes at the National Science Foundation, we institutionalized these interactions through the formation of a Research Group in Qualitative and Multi-Method Analysis. Each year, this authors' workshop convenes scholars for intensive discussion of one another's methodological research. By holding the meetings on the weekend between the week-long sessions of the regular institute, we have been able to double-dip on institute faculty, both as authors and discussants. Appendix 2 includes a list of projects discussed between 2009 and 2011.

The institute is now firmly established as the premier venue for graduate students and junior faculty to receive intensive instruction on the rapidly expanding qualitative and multi-method research canon. Moreover, the institute has had a direct influence on the new canon, especially in political science. There has also been a welcome expansion beyond political science (the institute's early focus) to attract an increasing number of students and faculty from cognate disciplines, including Sociology, Geography, Law, and Communications; and extending analytic coverage beyond pure qualitative methods—it now crosses several substantive boundaries (disciplinary and sub-field), and straddles epistemological cleavages. These developments are likely to continue as the institute becomes increasingly diverse.

One likely side-benefit of this increasing diversity is that the networking aspect of the institute, which is already very extensive, is likely to become even more effective. The most obvious advantage of attending the institute is that it provides training in qualitative/multi-method techniques and helps participants with their own research. A less patent benefit is that intensive interactions at the institute have established networks of participants who share methodological and substantive interests and who stay closely in touch, often for many years after the institute. As the institute continues to grow in size and diversity, there will be more opportunities for these kinds of interactions.

Let me close this brief essay where I began, by again applauding the incredible efforts of the instructors and participants at the ten institutes held to date. Their patience, stamina, and—above all—good humor have made the first ten years a remarkable experience. Let's hope that future iterations of the institute live up to this extraordinary precedent.

Appendix 1: Universities from which Centers, Departments, or Institutes were members of CQRM in one or more years, 2009–2011

Aarhus University, Denmark; American University; University of Arizona; Arizona State University; Bergen University, Norway; Boston University; University of British Columbia, Canada; Brown University; University of California, Berkeley; University of California, Irvine; University of California, Riverside; University of California, San Diego; Case Western Reserve University; University of Chicago; Columbia University; Cornell University; University of Delaware; Duke University; Emory University; University of Florida; Peace Research Institute, Germany; George Washington University; Georgetown University; Harvard University; Hong Kong University of Science and Technology, China; Indiana University; Johns Hopkins Baltimore; Johns Hopkins SAIS; Kent State; University of London, UK; Universite Laval, Canada; Leuven University, Belgium; University of Maryland; University of Massachusetts; Massachusetts Institute of Technology; McGill University, Canada; University of Miami; University of Michigan; Michigan State University; University of Minnesota; University of Montreal, Canada; New School for Social Research; Northwestern University; North Carolina-Chapel Hill; University of Notre Dame; Ohio State University; University of Oklahoma; University of Oregon; Oxford University, UK; University of Pennsylvania; Pennsylvania State University; University of Pittsburgh; Princeton University; Queen's University, Canada; Radboud University, Netherlands; Rutgers University; University of Southern California; University of Southern Denmark; Stanford University; Syracuse University; University of Sussex, UK; University of Texas; University of Toronto, Canada; Tufts University; University of Vienna, Austria; University of Virginia; University of Wales, UK; University of Washington; Washington State University; University of Western Michigan; Yale University.

Appendix 2: Projects Discussed at the Research Group on Qualitative and Multi-Method Research, 2009–2011 (year noted in parentheses)

Ariel Ahram, University of Oklahoma, “The Challenge of Conceptual Stretching in Mixed Method Research” (2009)
Karen Baird, Purchase College, SUNY, “Causal Explanations, Qualitative Research, and the Field of Public Policy” (2010)

- Jeff Checkel, Simon Fraser University, and Andrew Bennett, Georgetown University, "Process Tracing" (2010)
- David Dessler, College of William and Mary, "The Philosophical Architecture of International Relations: Theory and Narrative in the Study of World Politics" (2010)
- Thad Dunning, Yale University, *Natural Experiments in the Social Sciences* (forthcoming, Cambridge University Press), (2011)
- Tulia Falleti, University of Pennsylvania, "Toward a Comparative Sequential Method" (2009)
- John Gerring, Boston University, *Social Science Methodology: A Criterial Framework*, 2nd edition (forthcoming, Cambridge University Press), (2009)
- Adam Glynn, Harvard University, and John Gerring, Boston University, "Evidence of Causal Relationships: A Typology" (2010)
- Gary Goertz, University of Arizona and James Mahoney, Northwestern University, *A Tale of Two Cultures: Contrasting Qualitative and Quantitative Paradigms* (forthcoming, Princeton University Press), (2011)
- Jennifer Hadden, Cornell University, "Mixing it Up: Towards Synthetic Standards of Evaluation for Multi-Method Research" (2010)
- Diana Kapiszewski, University of California, Irvine; Lauren M. MacLean, Indiana University; and Benjamin L. Read, University of California, Santa Cruz, *Field Research in Political Science* (forthcoming, Cambridge University Press), (2009 and 2011)
- Kendra Koivu and Erin Kimball, Northwestern University, "Using SUN for Concept Formation and Understanding Causation" (2009)
- Kenra Koivu, Northwestern University and Erin Kimball, Northwestern University, "The Gap Within: Differences between Approaches in Qualitative Methods" (2010)
- Stephen Nelson, Cornell University, and Andrew Yeo, Catholic University of America, "Methodological Challenges and Progress in Ideational Research" (2009)
- Carsten Schneider, Central European University, *Set-Theoretic Methods: A User's Guide for Qualitative Comparative Analysis (QCA) and Fuzzy-Sets in the Social Sciences* (forthcoming, Cambridge University Press) (2010)
- Jason Seawright, Northwestern University, *Multi-Method Social Science: Combining Qualitative and Quantitative Tools* (forthcoming Cambridge University Press), (2011)
- Hillel David Soifer, Princeton University, "Permissive and Causal Conditions in Historical Causation: Windows of Opportunity and Types of Critical Junctures" (2009)
- David Waldner, University of Virginia, "Process Tracing, Its Promise, and Its Problems" (2011)
- Oliver Westerwinter, European University Institute, "Thick Network Analysis: Broadening the Methodological Repertoire for Studying Networks in World Politics" (2011)
- Sean L. Yom, Harvard University, "Theoretical Deduction or Empirical Induction? Resolving Causal Tensions in Comparative-Historical Research" (2009)

Notes

¹ This brief essay builds on a more formal treatment in Collier and Elman (2010). I thank Andy Bennett for his helpful comments on an earlier draft. The renaming of the Institute changed its acronym from IQRM to IQMR, but usage in all the essays has been standardized to IQMR.

² "Boot Camp" email, Elman to Bennett, January 28 2000. From 2002 through 2008 the institute was hosted by Arizona State University, and in 2009 it moved to Syracuse University.

³ This is the total number of instructor slots, not a count of different faculty, since some have returned to teach at different institutes.

⁴ We have from time to time had the occasional ad hoc arrangement where an organization could not join the Consortium but wished to support attendance by one of its employees, or where a scholar wished to pay for their own attendance. But these have been rare exceptions to the two traditional routes.

⁵ The NSF grants comprised: October 2002, Award No. 0350963, \$18,250; January 2005, Award No. 0452768, \$63,540; September 2007, Award No. 074968, \$7,582; June 2008, Award No. 0752589, \$103,386; September 2009, Award No. 0951774, \$1,572; and September 2011, Award No. 1124074, \$150,690.

⁶ The "how to" sessions on publishing and obtaining funding have typically been led by an editor of the *American Political Science Review*; and by a program director from the National Science Foundation's Political Science Program.

References

- Bennett, Andrew, Aharon Barth, and Kenneth R. Rutherford. 2003. "Do We Preach What We Practice? A Survey of Methods in Political Science Journals and Curricula." *PS: Political Science and Politics* 38:3 (July), 373–379.
- Collier, David and Colin Elman. 2010. "Qualitative and Multimethod Research: Organizations, Publication, and Reflections on Integration." In *The Oxford Handbook of Political Methodology*. Janet M. Box-Steffensmeier, Henry E. Brady, and David Collier, eds. (Oxford: Oxford University Press), 779–795.

Pam Bromley (IQMR 2003)

Pomona College

pamela.bromley@pomona.edu

I never would have imagined that participating in IQMR in 2003 would have fundamentally shaped my academic career, but it absolutely has. When I applied to IQMR, I hoped that the institute would shape the methodology for my dissertation. That the Institute was then held at Arizona State University in January and offered an escape from a cold Northeast winter certainly didn't hurt. What I learned at IQMR gave me the background I needed to complete my multi-method dissertation and to justify my very historical research to the many people in the discipline that are quantitatively focused.

But, in an unexpected way, IQMR has had a lasting impact on my career. Not all political science Ph.D. candidates are going to land tenure-track jobs; the job market is tough. And if you have a partner who is also an academic and you are trying to find two jobs in the same place, you might have to make some hard choices. That's what happened to me. I finished my Ph.D., a mixed-methods project in international relations, and followed my spouse and his tenure-track job across the country. I landed an adjunct position teaching writing, through an examination of globalization, at a nearby institution, and I began scouring the area for full-time positions. There weren't any tenure-track jobs in political science, but there was an open assistant director position at a local writing program. I played up my experience teaching writing as much as I could, which wasn't exactly hard: teaching political science basically demands that we teach students how to write about political science.

I got the job, and I credit IQMR with much of the success

I've had in the past three years. When a student wants to do an empirical independent study project in writing studies, I can direct them to the methodological readings I've found most helpful in setting up my own research projects. Because of my background in social science, my colleagues trust that I can set up, carry out, and report on multi-methods research. My coauthors and I are doing work that's never before been done in writing studies, fueled in large part by my background in methods. We're asking—and answering—important questions: whether the writing workshops we hold every semester for the biology department really help students write better lab reports; whether working consistently with the same writing tutor helps students improve their writing over the semester; whether conventional wisdom about what students get out of writing center visits really holds up. Of course, I still use what I learned at IQMR in my ongoing research in international relations. But I'm especially pleased to have brought what I learned at IQMR to composition studies, a field that has comparatively little experience with research methodology but is consciously seeking to move beyond anecdotal evidence in making claims.

Stephanie L. McNulty (IQMR 2003)

Franklin and Marshall College
smcnulty@fandm.edu

How did IQMR shape my scholarly trajectory? At the most basic level, IQMR made me think much more systematically about the methodologies that I use to research the questions that I ask. Before the institute, I had been exposed to statistical research methods and formal modeling—neither of which provided very much analytical leverage when thinking about local politics in Latin America. I had travelled frequently to South America to gather information in a, well, less-than-systematic way. Because qualitative method courses were not yet offered at my university, the moment I saw the advertisement—during the institute's second year—I (literally) begged my department to send me.

The institute gave me a new language to use when gathering, analyzing, and reporting my data. During my time there, some of the greatest minds in political science inspired me to think about preparing for the field, organizing information, conducting more systematic interviews, and carefully analyzing the causal linkages. My dissertation fieldwork, which is now published in my recent book, *Voice and Vote: Decentralization and Participation in Post-Fujimori Peru*, was much stronger as a result. More importantly, the skills that I gained through IQMR will continue to improve the stories that I tell in future scholarship.

The institute also taught me that we, as students of methodology, can contribute to these debates. I often think back to one day when two presenters, Deborah Yashar and Marc Morjé Howard, discussed their personal experiences doing fieldwork. They shared several of their techniques in the field, which I often use. Over the years I have come to realize that I, too, have a story to tell. As a result I am currently working on a monograph about conducting fieldwork with children in tow,

in hopes of helping other women juggle the demands of publishing and parenting early in our careers.

But the most important part of the institute is that it gave me a community in the field of political science. How many of us felt out of place at the traditional political science conferences before attending IQMR? Within days of arriving in Arizona, I realized that I did belong in this field and finally knew why. I met amazing scholars who introduced me to other amazing scholars, and all of our paths continue to cross. I do not know if I would feel at home in this field today had the institute not existed.

So, I want to thank the organizers of IQMR for helping us all become better political scientists. Not only has the institute improved our scholarship, but it has also given us a community. For this, I am truly grateful. Happy anniversary!

Sarita D. Jackson (IQMR 2004)

North Carolina A&T State University
sdjacks1@ncat.edu

The Institute for Qualitative and Multi-Method Research (IQMR) has had a tremendous impact on my research trajectory. As a doctoral student at Brown University, I attended the IQMR in the winter of 2004 at Arizona State University. The week-long series of workshops allowed me to develop a much deeper understanding of qualitative methods, their contribution to social science research, and appropriate research designs. My work continues to build off of the literature and notes on case-studies, path dependency, and conducting field research.

For example, a current article that has just been resubmitted to a peer-reviewed academic journal uses the information that I received while an IQMR student. This same article discusses why qualitative methods are best used in understanding the decision on the part of importers in some developing countries to use expensive inputs from developed countries; these decisions contradict widely accepted economic assumptions. Furthermore, the article pulls from some of the work of the IQMR faculty to highlight the benefits of a single-case study and the use of the path dependency framework to point to several non-economic factors that influence importers' decisions (in this case, textile importers from the Dominican Republic).

Additionally, the workshops on conducting field research still prove helpful. A series of articles on field research that appeared in the spring 2004 *Qualitative Methods* newsletter have also been useful. Since my work involves traveling to other countries and interviewing government officials and private sector representatives, it is imperative that I make the best use of my time in the field and gather appropriate yet substantial data. The information that I have received from the IQMR on field research was particularly useful when I traveled to the Dominican Republic in 2008 to conduct preliminary research and in 2010 with the support of the Fulbright Scholar Award.

My work in the Dominican Republic examined industry association participation in the DR-CAFTA negotiations from

the side room. Prior to traveling to the country, I made appropriate contacts with individuals from the media and statistical agencies who could provide background data that are not available on the internet. Before the actual trip, I also made contact with the government negotiators and representatives for specific industries being studied. Some interviews were conducted with embassy officials and other representatives based in Washington, D.C. The completion of these tasks prior to traveling to the Dominican Republic allowed me to conduct substantive interviews while in the country and even begin drafting articles. Therefore, if any additional questions arose, I could follow up with some of the interviewees or gather additional data while in the Dominican Republic.

In sum, the IQMR offered a better understanding of qualitative methods by providing students an opportunity to meet with and question scholars whose work is used in graduate research design courses. My work to date has made use of single case studies and the path dependency framework, both of which were discussed at length during my participation in IQMR. Finally, I continue to use the skills and methods pertaining to field research in my own field work.

Christopher Darnton (IQMR 2006)

The Catholic University of America
darnton@cua.edu

I attended IQMR in 2006, as a graduate student wrestling with a dissertation prospectus on interstate rivalry shortly before embarking on my fieldwork in South America. Five years later, I am teaching a graduate seminar on Qualitative Research Methods. Putting together the syllabus was both encouraging and provocative. It encouraged me to see how much important and useful material has emerged in the last half decade; it provoked, by uncovering how little work is still available on some seemingly essential topics and tools for graduate students who plan to write case studies. Although a wealth of new works clarify how to refine concepts, triangulate across multiple methods, and plan and execute fieldwork, for example, we seem to have comparatively little guidance on the critical analysis of primary and secondary texts (vital steps for most case studies), beyond Marc Trachtenberg's exemplary *Craft of International History*.

I would suggest that if qualitative methodology in political science has made important strides with regard to setting up theories and gathering raw data, more needs to be done to define standards, rules, and best practices for extracting reliable inferences from (or adequately explaining and understanding) the documents that confront us. As undergraduates progress from quizzes to research papers, and especially as graduate students pass comprehensive exams and embark on dissertations, the fundamental question is no longer "what do we think we know?" but "why do we think we know it?," since sources can easily be found for almost any argument. This is a central concern in my teaching because it's been a fundamental problem in my research. Pat explanations of seemingly overdetermined outcomes crumble when applied to relevant nega-

tive cases, while arguments drawn on research in one country or language weaken when sources from another are consulted, and similar degradation often occurs with the shift from secondary to primary, from current to historical, and from published to archival materials.

The centrality of these issues in my work is very much a consequence of my experience at IQMR, where presenters explained how changing their own research designs changed their findings and sometimes the conclusions of their fields, small workshops articulated tremendous variation in individual scholars' approaches to similar topics and questions, and collaborative conversations proposed alternative ways to frame and execute projects. This level of re-engineering might be comparatively easy for graduate students at the proposal stage, not yet confronted with the sunk costs of a long-term research project, but the IQMR spirit of constructive criticism on methodological grounds was inspiring, and it bears replication. As conference panel discussants, as peer reviewers, and as advisors, we might do well to model IQMR in miniature: exposing present flaws, identifying relevant alternatives, and encouraging constant improvement through sustained discussion and reformulation. Rewarding as it has been to catch up with members of the 2006 IQMR cohort at conferences and following their work, it is at least as reassuring to meet new colleagues and current graduate students, and discovering, partway through a discussion of an important emerging research project and potential ways to shore up its fortifications, that we bear the same methodological tribal markings. I take this to be the hallmark of IQMR: beyond instruction on specific tools, the cultivation of debate and inquiry.

Nivien Saleh (IQMR 2006)

University of Saint Thomas, Houston
Nivien.Saleh@gmail.com

In 2005, a few months into my first semester as a visiting professor of political science at Northern Arizona University, I learned that I would teach the department's largely qualitative Ph.D. course on research methods. To prepare me for this task, NAU agreed to fund my participation in the 2006 IQMR workshop.

Before coming to NAU, I had been a Ph.D. student at American University's Department of Government. Like many other doctoral programs in the United States, ours placed heavy emphasis on quantitative methods and treated qualitative inquiry as an inferior substitute. Having served for two years as the teaching assistant for statistics, I understood its logic of inference well but disagreed with it. What captured my imagination was critical realism, an alternative methodological paradigm about which I knew little. Thus, when NAU enrolled me in the IQMR workshop, I perceived this as a chance to put my affection for realism on a firmer footing.

IQMR turned out to be one of the most intellectually satisfying experiences of my life. Many of the instructors had just published books on the methods of inquiry. They were passionate about their subjects, eager to explain, and happy to forge bonds with a new cohort of social scientists. I seized the

precious opportunity to interact with these methodologists, quizzing them on the strengths and weaknesses that I perceived in their arguments.

The Ph.D. course on methods that I later taught at NAU clearly benefited from this experience. Meanwhile, my personal hopes of learning critical realism remained unfulfilled, for while the distinguished professors who taught at IQMR tried to break with the hold that statistical inquiry had over political science, they remained committed to some of its basic assumptions. This disappointed me at first, until I finally understood that I was no longer a student but a fully credentialed member of the discipline. As such, it was up to me to contribute to the debate and steer it in the direction I thought appropriate. This insight plus many, many hours of work led to the article “Philosophical Pitfalls: The Methods Debate in American Political Science,” which the *Journal of Integrated Social Sciences* published in 2009. Unlike the more established disciplinary journals, it agreed to make its content freely available on the web, which was important to me.

The article holds that qualitative inquiry’s poor ability to draw causal inferences from examined to unexamined cases results from its commitment to neopositivist foundational assumptions. Within critical realism, the article argues, causal inference from even a single case is possible. My book *Third World Citizens and the Information Technology Revolution*, which Palgrave Macmillan published in late 2010, applies this claim. Even though I am proud of the study’s substantive findings on technology diffusion, I believe that the book’s greatest intellectual accomplishment is its depiction of a new approach to causal inference.

IQMR has been an important part of my learning process. I could not have arrived at the insights I generated without the guidance of my workshop mentors. Their ideas inspired my own thinking and enabled it to develop from a half-baked intuition into a fully formed argument. In my view, this is what scholarly community accomplishes when it is at its best.

Halit Mustafa Tagma (IQMR 2007)
Sabanci University, Turkey
mustafatagma@gmail.com

Attending the annual Institute for Qualitative and Multi-Method Research (IQMR) has been one of the most beneficial experiences in my academic career. Prior to my attendance in 2007, I had taken two courses on qualitative methods at Arizona State University and the institute further deepened my exposure to multi-methods. We were exposed to various methodological issues and stakes in our discipline. I was able to not only expand my methodological skills, but also extend my professional network, which to this day provides me with various research opportunities. My exposure to multi-methods in political science has allowed me to publish on various research puzzles using different theoretical approaches. Let me briefly go over some areas where qualitative methods and the institute have directly influenced my research trajectory.

Case studies and process-tracing have been central to my research. A few years ago, I was trying to explain why Turkey

has persistently desired to be a part of the European Union even though there were so many obstacles in front of it. The answer that I, along with my co-author, came up with was from a historical institutionalist perspective. This approach required one to be able to show the negative feedback mechanisms and positive incentives that institutions impose on actors. Our argument was that following a critical juncture, Turkish modernizing elite followed a path-dependent foreign policy orientation that was pro-European. Process tracing the logic of path dependency in our case, we were able to link institutions and historical legacies to current political preferences of policy-makers.

In another publication, I benefited from Charles Ragin’s take on the comparative method and Colin Elman’s illustration of property-space during the IQMR meetings. Early during my career, I was trying to explain why states would voluntarily forego nuclear weapons as this appeared to be an anomaly for the Realist theory of international relations. The feedback that I received during my presentation in front of the IQMR audience led me to concentrate on possible necessary and sufficient conditions from a comparative perspective. It was in part the concepts and methods that I was exposed to at the institute that enabled me to isolate the necessary variables to solve this puzzle from a Realist perspective.

The lectures on research design at the annual institute were golden opportunities for those of us seeking external funding. Many of my grant applications were crafted with key points and ideas that were brought up during the session on how to construct a research design and write grant proposals. It is part due to this experience with methodological skills that I was recently awarded with a generous Marie Curie grant by the European Commission for research on nuclear proliferation. The referees for my project had praised my grant proposal’s epistemological and methodological awareness of the issues and stakes in carrying out qualitative research.

My research on critical and constructivist perspectives on the “war on terror” has also benefited from the lectures at IQMR. Aided by the lecture on how to conduct archive-studies by James Goldgeier, the participants of IQMR were fortunate to learn hands-on research techniques on accessing, analyzing, and interpreting archives to aid their research. This has been particularly useful for those of us who have employed archive work in political science. Ted Hopf and Lisa Weeden, who have given lectures and advice on using appropriate data-collection techniques, have provided the methodological backbone to my research in this area. The openness of IQMR members and participants to multiple methods and various epistemological perspectives has greatly influenced my research trajectory.

Finally, my experience with the consortium has given me the background to offer a qualitative research methods course in my university. This course is going to be one of the first courses in Turkey that illustrates qualitative research methods in political science, and my syllabus has been influenced by the lectures and sessions that I have attended. The syllabus database, made available on the institute’s website, has been particularly helpful for those of us constructing reading lists. I

thank all who came up with the idea of the database, and encourage other members to share their syllabi as well.

All in all, the consortium has contributed to major dimensions of my academic career, including publications, teaching courses, obtaining grants, and expanding my research network. It is because of the exposure I had at the IQMR to multi-method research that I have been able to publish in the various sub-fields of our discipline. Over the years the prestige of the consortium has grown tremendously, and I have personally come to see its intellectual influence in several continents. Once again, my thanks go to all the administrators, scholars, and participants who have made this wonderful experience possible.

Maria Koinova (IQMR 2007)

University of Amsterdam, The Netherlands
m.v.koinova@uva.nl

Congratulations to IQMR on its ten-year anniversary! I am deeply grateful to all the professors from the 2007 winter school in Arizona, who opened the doors for me to better understand the links between qualitative and quantitative methodology. The IQMR experience was a trigger for a learning process about multi-methods research, which recently helped me to win a 1.498.000 EUR institutional grant from the European Research Council (ERC). During the next five years I will build a research group to study the transnational diaspora mobilization in Europe and its impact on political processes in the Balkans, the Caucasus, and the Middle East.

As a post-doctoral researcher at Cornell in 2007 I was just starting my new research program on diasporas and contested sovereignty. IQMR's session aimed at giving personalized feedback on draft proposals was of great benefit to me. I want to thank especially Andrew Bennett for his helpful comments at that initial stage.

IQMR has a unique atmosphere of openness towards different methodologies, excitement about exploring new frontiers, and top-level expertise. This inspired me to further follow a short course of the Qualitative and Multi-Methods Research section at APSA in 2010 and different APSA panels. For my ERC project I did significant additional research, but the IQMR experience was defining.

Within the ERC grant my research group (I as a PI, a post-doc, and two Ph.D. students) will study on a comparative basis how the Albanian, Armenian, Bosnian, Iraqi, Kurdish, and Palestinian diasporas mobilize in five European states—France, Germany, the Netherlands, Sweden, and the United Kingdom—in support of political processes taking place in their original homelands. The project considers also diaspora mobilization on the supra-national level, including the European Union, the United Nations, and other international organizations. This project utilizes sequencing of qualitative and quantitative methodologies. We start with comparative studies and then test the developed theories with a cross-country representative survey to be conducted among 25 country-groups in five EU states. The survey will establish a much-needed quantitative

dataset sensitive to both transnationalist processes and specific countries.

I joined IQMR in 2007 as an “open-pool” candidate. I would like to encourage the organizers of this school to open in the future more slots for scholars from different parts of the world, who would be able to join and gain such a valuable professional experience.

Happy anniversary!

Lily Hsueh (IQMR 2009)

University of Washington
lhsueh@u.washington.edu

As an economist who had left the methodological homogeneity of the economics discipline for the interdisciplinary field of environmental governance and policy, my instinct was that research designs that encompass both qualitative and quantitative approaches are the key to understanding the complexity of transboundary and multi-scale environmental problems, such as climate change and toxic waste governance. While I had the foresight that the explanatory power of any research endeavor is enhanced when multiple methods are appropriately employed together, I did not know the epistemology or the mechanics of executing mixed methods. The Institute for Qualitative and Multi-Method Research (IQMR) became the training ground where I gained knowledge of the philosophical underpinnings of qualitative methods and the complementarities of alternative approaches. IQMR has influenced my research trajectory in the nature of the questions I ask and the iterative critiques in which I engage in regarding the uses, strengths, and limitations of the methods I employ.

One of the main lessons I have drawn from my training at IQMR is that my research inquiries need not be limited to the methods I know and certainly not to the methods that are in vogue. I am no longer afraid to ask research questions about entrenched regulatory institutions and authorities that require primary data collection in the form of in-depth fieldwork through survey-guided or open-ended interviews because quantitative data are not available or are unsuited for explicating institutional structures and policy processes. I have learned that letting the research questions drive the methods I use and being open to diverse techniques for collecting and analyzing data are the hallmarks of a scholar seeking objective truth.

I am currently pursuing a Ph.D. in environmental economics, politics, and policy at the University of Washington. In my dissertation, I utilize a mix of quantitative and qualitative methods to examine the ways in which regulators, firms, and NGOs engage one another in the marketplace and the regulatory regime to address pollution control. In these variations of private authority, firms negotiate with regulators to curtail the use of hazardous chemicals in their production process.

In my quest for causal inference, IQMR has inspired me to engage in iterative critiques of the methods I use. My foray with Bayesian statistical methods is an example of this: In the context of model uncertainty where the true model is unknown and multiple competing specifications exist, Bayesian Model

Averaging techniques offer causal leverage. Structural break analysis from both the frequentist and Bayesian traditions account for “big events” like wars, oil price shocks, financial crises, policy interventions, and technological innovations.

While answering important causal impact questions, these quantitative approaches fall short of uncovering the causal mechanisms at work. As such, I have turned to comparative case studies whereby in-depth, open-ended interviews with purposively selected informants allow me to trace plausible causal processes for building a framework of compliance and enforcement under which regulators and regulatees negotiate to identify pollution sources and means to address them. IQMR has certainly influenced the methodological course of my academic career.

Juan A. Bogliaccini (IQMR 2009)

University of North Carolina, Chapel Hill
jbogliac@unc.edu

The IQMR Institute was a great experience as a graduate student for many reasons, but in terms of my research trajectory it influenced me in three main aspects. First, it fostered profound reflection on the epistemological foundation of my research in political science, particularly regarding the ontology of different logics of inquiry. Upon concluding that these logics shared an ontological principle, I became interested in exploring nested research designs that combined both quantitative and qualitative techniques.

Second, the IQMR Institute reinforced my conviction regarding the academic relevance of exploring the theoretical links between causes and effects, which are usually not easily supported by observations. This quest for explanatory links is only fruitful, in my opinion, if the researcher is able to successfully describe the phenomena under study. In this sense, IQMR called my attention to the importance of “descriptions” in political science, something that has been relegated to a secondary place given the limited academic rewards for such kind of inquiry. However, without good descriptions, there are few good explanations.

Third, the Institute provided me with useful tools for qualitative and “small N” quantitative research. In particular, I rediscovered the importance of archival research for qualitative social science and utilized it heavily in my research on bargaining over redistribution in post-neoliberal Latin America. In particular, I conducted extensive research on Parliamentary Sessions and Commissions’ documents and press reports that yielded important information on how labor, business and political parties bargained over laws on industrial relations and taxes. The combination of this research technique with interviews allowed me to answer my substantive questions as well as the virtues and limitations of using interviews for social research.

To sum up, IQMR served as a landmark in my graduate trajectory. Being a comparative political science major with a minor in methods, I found that IQMR provided an excellent opportunity to think critically on both the discipline and my

own conceptual and methodological approaches to political phenomena. On a side note, it also provided a great opportunity to develop networks with other graduate students and scholars interested in substantive and methodological puzzles similar to mine. Overall, the IQMR is a great and worthwhile experience.

Catherine Herrold (IQMR 2009)

Duke University
ceh21@duke.edu

I arrived at the IQMR after two years in Duke University’s Public Policy Ph.D. program. At that time I had a very vague notion that I wanted to study the role of foundation-based philanthropy in the Arab region, but I was unsure how to plan, structure, and implement such a project, much less how to begin to narrow the topic in the first place. After IQMR, I not only had a much more clear sense of how to approach a qualitative research project; I had also developed a network of political science colleagues with whom I would convene conference panels, share experiences in the field, and maintain strong friendships.

The classes at IQMR introduced me to important concepts that I used to develop and conduct my case study research, while the feedback that I received on my project proposal encouraged me to both narrow my question and consider new literatures and theories to incorporate. The session that focused on conducting fieldwork proved to be particularly useful, ensuring that I arrived in the field with the necessary organization tools and strategies for soliciting and conducting interviews.

In recent years I have collaborated with a number of IQMR colleagues to convene conference panels focused on civil society in the Middle East. We have continued to learn from each other in the process, exploring ideas, honing theories, and finding new areas for research. I have also benefitted from meeting IQMR colleagues in the field. They have not only offered advice on conducting research in the local context, but have also been an important source of friendship and support during what can sometimes be a lonely process.

I am grateful for the knowledge and ongoing support that I have gained from IQMR, and I hope that many future political science students will be as lucky as I to be a part of the IQMR community.

Steven Samford (IQMR 2009)

University of New Mexico
ssamford@unm.edu

This past year I went door-to-door through the streets of dusty Mexican villages looking for family-run ceramics workshops to survey. Over the course of my dissertation fieldwork, particularly when the task seemed most daunting, I found myself reflecting on how I arrived at that point. Throughout my doctoral studies I had become increasingly interested in economic development policy and its impact on small businesses,

Kate Svyatets (IQMR 2009)
University of Southern California
svyatets@usc.edu

but I had never foreseen doing fieldwork that felt so concrete and so particular. Partly responsible for this micro-level approach was my experience at IQMR, which helped shift not the consequence of the theoretical and empirical issues I was interested in addressing but the scope of the data that I felt appropriate to address them.

I attended IQMR as I was in the process of refining a dissertation topic that grappled with the claim made by endogenous growth economists that technological change is fundamental to economic growth, pushing firms' capacities to adopt and governments' capacities to promote (or inhibit) adoption to the forefront. Perhaps the most significant manner in which the Institute affected me was in helping divest me of the notion that I needed to have "big" data in order to address this issue or that a broad, cross-national approach was necessarily the best way to approach a study about the adoption of technologies. After all, while on an aggregate level the large-scale adoption of improved technologies by firms could produce notable development outcomes for a country, technology is ultimately adopted by individual firms or workshops, and, in many labor-intensive sectors, technological advancement is the exception rather than the rule. To have an understanding of why small firms do not adopt available technologies, I needed to approach the question on a micro-scale: What informational or financial barriers do they face individually, and how exactly does policy make those surmountable? IQMR lectures on case-selection and combining qualitative and quantitative research shaped the logic by which I selected locations for surveying workshops. Courses on ethnography and fieldwork improved my understanding about how I might go about gathering workshop-level data in a developing country, and prepared me for some of the challenges that I indeed faced. Perhaps less tangibly, the Institute also helped me justify my use of these methods and a finer-grained approach to the issue of technology adoption and economic development. Moreover, it gave me confidence that there is a place in the discipline for scholars like me (perhaps like "us," if you are reading this) and that work of a more qualitative bent would find an audience among my colleagues.

By my current understanding, the capacity to adopt new technologies depends upon factors ranging from government engagement to network connections between firms (formal and otherwise). Coming to this understanding required that I be in the field and in workshops, that I conduct in-depth interviews, observe meetings of producers, survey producers formally, and ultimately combine the information gathered from all of these sources. While not at all the "big" data I imagined I would compile, what I can say about it is that I have confidence in the causal story that has come from them, a story about the microfoundations of adoption of new technologies by small firms in labor-intensive industries.

I attended IQMR in 2009, when I was starting to write my dissertation, and IQMR was a very useful experience for me. I am very grateful to Colin Elman for organizing the summer school at the Syracuse University. It was extremely helpful for shaping the methodology part of my dissertation and for networking with colleagues from other universities.

It was my great pleasure to learn firsthand from such prominent scholars as David Collier, Jason Seawright, Gary King, Andrew Bennett, James Mahoney, Gary Goertz, James Goldgeier, Elizabeth Saunders, Ron Rogowski, Thad Dunning, Sid Tarrow, Nina Tannenwald, and others. Before IQMR, I had read their work in classes and research, and I was very excited to meet them in person and to attend their seminars and lectures at the school.

The module that was especially relevant to my research design and, therefore, extremely helpful to me, was "Designing and Conducting Field Research" by Diana Kapiszewski, Julia Lynch, and Lauren Morris MacLean. In addition to an excellent session, they gave us a very detailed handout that I still keep and refer to from time to time.

Another aspect of the school was presenting our dissertation proposals and getting feedback from colleagues in the class. This was a great chance for me to present my research, to articulate my research goals, and to see which areas of my proposal needed clarification and improvement.

In addition to the learning experience, the IQMR summer school provided great opportunities for networking. I met colleagues from all over the United States, some of whom became my good friends. In my view, this aspect of IQMR is as important as the knowledge of methods that we acquired during the summer school. I am also grateful to Colin Elman and other professors who encouraged us to meet new people in the classroom rather than just socialize with friends from our home universities. Such tips in the classroom made a difference.

Finally, I would like to thank IQMR for excellent logistics and living conditions that we had in Syracuse. For the duration of the school, we stayed at a Sheraton hotel, within walking distance of our classes. We also had full access to the university facilities, such as libraries and the gym. IQMR always provided us with coffee and snacks during classes, which was great, as we did not have to think about such things and run around campus looking for Starbucks.

The IQMR summer school was a great learning and networking experience, and I highly recommend it to all Ph.D. students. I am grateful that my program, Political Science and International Relations at the University of Southern California, provided me with an opportunity to attend IQMR.

Ajay Verghese (IQMR 2009)
The George Washington University
ajay@gwmail.gwu.edu

For about my first three years of graduate school I talked rather opaquely about doing “some kind of statistical analysis” in my dissertation project. I was hazy on the details, of course, and had no real idea about *why* I wanted to do it except that you were *supposed* to do it. But to be honest, I didn’t really want to do any kind of statistical analysis. I just wanted to hunt around in faraway archives. And I wanted to go somewhere where people would tell me that was totally a great idea. So I attended IQMR in the summer of 2009 in Syracuse. But, to my surprise, I ended up learning just as much about quantitative methods as I did qualitative methods. And my dissertation now includes a statistical analysis after all. And I think it’s a pretty decent one. How could this have happened?

First, at IQMR I did learn to embrace my inner historian. After years of subtly and not-so-subtly being taught that qualitative methods deserved a lesser position to quantitative methods, I finally figured out that it doesn’t have to be this way. The message may seem like a truism, but in my discussion with other graduate students it seems like this lesson takes some time to stick. Just glancing over the roster of instructors opens your eyes to the wealth of impressive scholars who probably don’t even own Stata! So why was I so concerned and sheepish about doing archival research? IQMR taught me to embrace my strengths. Why be a second-rate mathematician when I could be a pretty good historically-oriented political scientist? And why use a methodology that didn’t really interest me in the first place? I realized that I’d end up writing a much better dissertation in the long run if I was passionate about how I was going about it.

But then something strange also happened at IQMR. I ended up deciding that I did want to include a statistical analysis. I had the data. Why not learn how to analyze them? And I looked around and noticed that these IQMR people weren’t Luddites, rebelling against fancy equations they didn’t really understand. Rather, they simply refused to have an uncritical acceptance about statistical methods. And that’s a good thing. But you could go one step further—one of the lessons I learned was that if anything, quantitative scholars don’t learn about the *wide range* of statistical techniques out there. They use OLS and call it a day. But what about matching designs? Or an instrumental variable analysis? I learned at IQMR that we young scholars have it pretty good—we have a wide array of methods, and we’re learning every day better and better ways to put them together.

If multi-methods are here to stay in political science, you’d be making a big mistake not attending IQMR. There’s no better place to learn about how to put together qualitative and quantitative methods (and also when not to!). And you won’t find a preference for either side. Simply put, you’ll learn how we can all get along.

Kerry Crawford (IQMR 2010)
The George Washington University
kfc@gwmail.gwu.edu

Like most members of my cohort of graduate students I grew up in a world in which graphing calculators and computers were almost permanent extensions of one’s person. My high school statistics textbook looked more like a TI-86 instruction manual than a math textbook. My undergraduate and graduate courses in Political Science required a good bit of dabbling in STATA and R. The majority of this current generation of graduate students, and certainly future generations, have lived in a world in which impressive computer programs were thought to be able to answer virtually any question or satisfy any curiosity.

Despite the tech-savvy world in which I have come of age, my hunch has always been that there are a great many big and important questions that cannot be answered in full by quantitative methods alone. I do not discount the value of quantitative methods for political science; however, my research delves into complicated, messy, and utterly human issues, asking questions and searching for answers that cannot be stuffed into clean models and regressions. When I decided to write my dissertation on the ways in which states and international organizations respond to wars involving widespread sexual violence against civilians, I knew I would need to learn as much as possible about working in archives and conducting interviews.

I attended the Institute for Qualitative and Multi-Method Research (IQMR) after completing the second year of my doctoral program. I knew that I wanted to finish and defend a dissertation prospectus by the end of my third year but I also sensed that my limited experience with and training in qualitative methods would hamper my progress. My firm belief, and one that I know is shared by many of my peers, is that all graduate students could benefit from greater qualitative and mixed-methods training in addition to the quantitative training emphasized in the early years of graduate education.

Attending IQMR did more than provide me with the tools I needed in order to conduct an effective interview and find relevant documents in the archives: The two weeks in Syracuse validated my assumption that I was not the only political science scholar dissatisfied with the field’s increasing pressure to employ quantitative methods. IQMR allowed me to engage with other graduate students struggling with methodological angst similar to my own. The sessions also gave participants the valuable opportunity to learn from and interact with junior and senior faculty members from universities across the country who are not only utilizing qualitative and mixed methods in their research, but are doing it well and building successful careers out of it.

The methods I learned about during IQMR have, of course, helped me in the construction and execution of my dissertation research plan; yet, what has had the greatest influence on my research is the sense that there is a large and vibrant community of scholars using qualitative and mixed-methods to

conduct rigorous and interesting research. The answer to my research question has to come from six decades of archived documents and the keen insights of the very people who have worked to put a women's human rights issue on the international agenda; it requires a great deal of investigative research. Participating in IQMR underscored my belief that "rigorous" does not have to mean "full of regressions" and that scholars can and should continue to ask the big, important, messy questions because we do, in fact, have the tools to answer them.

Anastasia Shesterinina (IQMR 2010)
University of British Columbia
ashester@interchange.ubc.ca

"Do not select on the dependent variable. Do not advance no variance designs. Do not study single cases... And do not even think of interpretivism!" A young graduate student may receive these "do nots" early on in her studies as a mantra to follow in her political science career. Entering a discipline for decades dominated by epistemological commitment to positivism and quantitative methodological approaches leads to a number of conclusions. Maybe the "do nots" are there for a reason and accepting them is not such a bad idea after all.

Memorizing the mantra, however, does not answer many questions. What if a researcher is concerned with a particular outcome? What if she cares about a particular case or set of cases important to that outcome? And what if positivism is not fully convincing to her, regardless of its status in the discipline? In a word, how to productively go counter to the general flow, following one's intuitions, and defend oneself against the discipline's "do nots"?

Having arrived at Syracuse with a "split personality"—a set of correlational hypotheses and an urge, driven by the need for success in the discipline, to estimate average effects, on the one hand, and a causation-oriented, single-case project drawn from a process-oriented mind, on the other—I quickly learned that my "faults" were not just my own. I was surrounded by others with similar questions, a group of IQMRers that now comprises the network of my methodology go-tos.

Apart from the ongoing support of this vast cohort of young researchers in political science, the Institute provided me with powerful tools and feedback regarding my specific interests in the discipline. The knowledge and advice of the experts brought together by the Institute has guided my search for a research niche in political science and significantly shaped my Ph.D. dissertation trajectory.

The insights shared at the IQMR by David Collier and Charles Ragin on the trade-offs between case- and variable-oriented research offered me a broader picture to ground my choices in. Jason Seawright's advice on mixed methods helped me develop a research design suitable to the population of post-conflict cases that I am interested in—a mix of dataset construction, case selection, and in-depth within-case analysis placed in a regional and cross-national comparative perspective, all aimed at increasing the certainty of my findings. I used Gary Goertz' guidelines to construct the concept of organized collective political violence, the heart of my dissertation.

I followed Colin Elman's suggestions on conjunctural causation and Andrew Bennett's hints on process tracing to approach the complexity of the problem of post-conflict violence that I address. Lisa Wedeen's wisdom on close engagement and interpretation helped me prepare to venture into fieldwork in conflict areas such as Georgia.

Against all "do nots," from a confused graduate student entering the discipline emerged a methodologically well-equipped young researcher able to identify the benefits and shortcomings of using one or another approach, defend her choices, and when asked, present an array of tools used in designing her research.

Patrick Mello (IQMR 2011)
Humboldt University, Berlin
patrick.mello@sowi.hu-berlin.de

The Institute for Qualitative and Multi-Method Research (IQMR) is truly unique in a number of ways. First, unlike other summer schools or workshops, most of which focus on teaching specific methods, IQMR takes a broader approach and introduces participants to a wide range of methodological topics over the course of two weeks—from concept formation to typological theorizing, fuzzy-set analysis, counterfactual models, and natural experiments, to name just a few of the modules offered in 2011. Each course is taught by a leading scholar on the subject, which makes for an excellent opportunity to deepen one's knowledge of a particular topic or to get introduced to new methods and approaches. Second, in addition to the substantive courses, participants get to present their own research projects and discuss their design in a smaller setting with colleagues and faculty members. I found these research design discussion groups to be one of the greatest strengths of IQMR. Due to the cooperative spirit of these sessions each project received a fair amount of constructive critique and plenty of helpful comments—way beyond what one usually encounters when presenting a paper at a conference—and often times the discussions continued over lunch and throughout the breaks.

How did IQMR affect my own research trajectory? For one thing, I found it tempting to conceive of alternative ways to approach my research question, which investigates the observed variance in Western democracies' participation in coalition warfare in the past two decades. Is it possible to think of a persuasive natural experiment given this research aim? How to combine a comparative design based on fuzzy-set Qualitative Comparative Analysis (fsQCA) with either process tracing or quantitative approaches? Indeed, many of the questions I encountered throughout the institute evolved around combining different methods without losing inferential leverage or diluting my research aim. Overall, I believe that IQMR further raised my methodological awareness regarding the strengths and drawbacks of specific methods and approaches. In addition, I learned about several methods that I would love to apply in future research projects. As a European researcher I also found it intriguing to get first-hand exposure to ongoing methodological debates among my North American colleagues.

It was refreshing to see that a number of topics remain controversially discussed and that real disagreements exist concerning the respective merits of certain methods and approaches—even among those who strongly support the notion of methodological pluralism. I am convinced that it is important to spell out these methodological perspectives, which are often

linked to ontological and epistemological differences. Fortunately, the institute provided a great sounding board for these different views to engage in a critical dialogue. On a final note I would like to thank the organizers and people behind IQMR for all their work and the great opportunity they are providing!

Symposium: Methodological Issues in Medium-N Research

Between a Rock and a Hard Place: The Uneasy Position of Medium-N Research in Comparative Political Science

Patrick Emmenegger

University of Southern Denmark
emm@sam.sdu.dk

In their seminal “tale of two cultures,” Mahoney and Goertz (2006: 227) argue that “quantitative and qualitative research traditions can be thought of as distinct cultures marked by different values, beliefs, and norms.” They go on to contrast these two different research traditions across ten areas. Although their discussion of the qualitative research tradition is partly based on the work of Ragin (1987, 2000), it becomes quite clear that they mostly draw their inspiration from small-N research, on the one hand, and (quantitative) large-N research, on the other. Research based on the analysis of an intermediate-sized N (say, around 20 to 30 cases) does not play a prominent role in their account. For instance, typical medium-N research designs are not oriented towards positive cases on the dependent variable (as most qualitative research). Nor can the cases be randomly selected on the independent variables (as it is good practice in quantitative research). Rather, in medium-N research designs the cases analyzed correspond to the total population of cases. Hence, medium-N research poses unique methodological challenges.

Research based on the analysis of an intermediate-sized N, typical of much welfare state and comparative political economy research, is often caught between a rock and hard place. Econometrical methods (the quantitative tradition) are of limited use to gauge the existence of empirical regularities because the number of cases is not sufficient for appropriate model specification. Facing the methodological challenges of medium-N populations, many authors increase the number of observations by using multiple observations per country in a pooled time-series cross-sectional design or extend the number of cases by softening the scope conditions. However, both strategies are problematic. Pooled time-series cross-sectional designs exchange the medium-N problem for new econometrical problems and have no answer to the fundamental limitation that most relevant information is often cross-sectional. Increasing the number of cases by softening the scope conditions is likely to result in concept stretching. In addition, it entails the risk of ignoring differences-in-kind in the pursuit of explaining differences-in-degree. Alternatively, researchers can circum-

vent the medium-N problem by focusing on a small selection of cases (the qualitative tradition). This, however, raises the ever-looming problem of case selection and questions regarding external validity. What can we say about the whole population of cases, the main goal of medium-N research, when we cover only a handful in our actual research?

One option could be to abandon medium-N research in favor of more large-N and small-N research respectively. However, there are simply too many medium-sized populations. For instance, the European Union has 27 member states, the Organisation of Economic Co-operation and Development 30 and the Organization of American States 35. The African Union has 53 member states, while the Arab League has 22. The collapse of the USSR led to 28 post-communist countries. On the substate level, we have 16 Länder in Germany, 26 cantons in Switzerland, and 50 states in the USA. The People’s Republic of China has 22 provinces and claims, but does not control, a 23rd one. In 2010, 22 countries have been using the Euro as their main currency, even though only 16 countries are represented in the European Central Bank. As of 2010, 32 states have or have had a female head of state. Wherever we look, we encounter medium-N populations. If political science is understood as a problem-driven discipline, then we need a methodology for medium-sized populations.

This symposium brings together six essays on issues and problems typical of medium-N research. Ebbinghaus and Kvist highlight the unique methodological challenges of medium-N research and discuss possible remedies. The contributions by Goertz and Kenworthy deal with pooled time-series cross-sectional designs and provide guidance on when these designs can be useful. Emmenegger and Petersen examine how area experts perceive state-of-the-art comparative political economy work. Finally, the contribution by Shalev offers suggestions on how to overcome potential problems of medium-N research and even turn them into strengths. Together, these contributions demonstrate that medium-N research designs—despite their uneasy position in social science methodology—should be a core part of our methodological toolkit.

References

- Mahoney, James and Gary Goertz. 2006. “A Tale of Two Cultures: Contrasting Quantitative and Qualitative Research.” *Political Analysis* 14:3, 227–249.
- Ragin, Charles C. 1987. *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*. Berkeley: University of California Press.
- Ragin, Charles C. 2000. *Fuzzy-Set Social Science*. Chicago: University of Chicago Press.

Case Selection in Medium-N Comparative Welfare State Analysis

Bernhard Ebbinghaus

University of Mannheim, Germany
ebbinghaus@uni-mannheim.de

Comparative welfare state research has been relying on medium-N analyses to test macro-typologies of institutional configurations. Most prominently, Gøsta Esping-Andersen's *Three Worlds of Welfare Capitalism* (1990) led to a plethora of research on the merits of fitting real welfare states across Europe and beyond into his three ideal-type categories. A "regime" approach is useful in comparative analyses of welfare states in order to conceptualize distinct typologies in which to classify empirical similarities and differences (Lange and Meadwell 1991). While ideal-type regimes should be theoretically grounded, comparative typologies commonly seek to classify empirical real cases, sometimes rather inductively. Since advanced welfare states are not a worldwide phenomenon, comparative studies cover a deliberate selection of European, North American, and other OECD countries. These quantitative analyses commonly apply cluster analysis, linear regression, or pooled time-series of medium-N cases. Qualitative Comparative Analysis (QCA), advanced by Charles Ragin (1987), has become a mixed-method tool to analyze cross-case configurational typologies, often building upon in-depth within-case studies of particular welfare states. This essay will discuss the methodological problems of selecting cases in medium-N research designs, using quantitative or qualitative comparative analyses of welfare state regimes as examples.

Comparative institutional analyses on welfare states go beyond the classification of particular institutions; more ambitiously, they aim to capture cross-national variations in "regimes." Most prominently, Esping-Andersen uses his welfare (state) regime approach "to denote the fact that in the relation between state and economy a complex of legal and organizational features are systematically interwoven" (Esping-Andersen 1990: 2). Similarly, the Varieties of Capitalism approach (Hall and Soskice 2001) stresses the "systemness" of institutional arrangements leading to different types of market economies, partially mirroring welfare regimes (Ebbinghaus and Manow 2001). For such institutional regime analysis, typologies have "an obvious attraction in being able to characterize whole systems with the related implication that different systemic features hang together" (Lange and Meadwell 1991: 84). However, as Esping-Andersen concedes, the welfare regime approach implies a trade-off: "Since our intention is to understand the 'big picture,' we shall not be able to dwell on the detailed characteristics of the various social programs" (1990: 2).

In my essay I will discuss several methodological questions concerning case selection in comparative cross-national analyses. First, based on Esping-Andersen's threefold regime typology (1990; 1999) quantitative studies of advanced wel-

fare states will be discussed in respect to the geographical selection bias. In the second part, I will discuss whether quantitative medium-N analysis is biased through the selection of cases restricted to member countries of OECD and/or EU. Finally I will address the question whether Qualitative Comparative Analysis (QCA) and its Fuzzy Set (fs) extension are more appropriate to deal with the selection problems of medium-N comparative analysis and are more suitable for analyzing typologies. My main conclusion is that we should be careful in generalizing from quantitative medium-N analyses without due consideration of selection problems in cross-national analysis. Instead it might be more appropriate to explore welfare regime variations with the help of QCA or fsQCA, thereby explicitly discussing the rationale in selecting cases and their specific configurations.

Country Selection and Welfare State Typologies

Esping-Andersen's seminal regime typology was based on a comparative historical analysis of social policy development in 18 OECD countries up to the 1980s (Esping-Andersen 1990), going beyond linear models of social expenditure of medium-N cases. The key dimensions of Esping-Andersen's typology are the degree of de commodification, that is, "the degree to which individuals, or families, can uphold a socially acceptable standard of living independently of market participation" (1990: 37), and the kind of social stratification fostered by social policies. The three distinct regimes refer to political movements and their ideas generating different welfare models: the Liberal (Anglophone countries), the Conservative/Corporatist (continental Europe and Japan), and the Social Democratic (Scandinavia) regimes.

The ideal-typical distinction of the three welfare regimes has been subject to various reconsiderations over the past decades, some scholars arguing in favor of four or more regime types (see Arts and Gelissen 2002, 2010). Although Esping-Andersen's quantitative analysis included 18 OECD countries, it did not include the new Southern EU members of the 1980s: late-democratizing Greece, Portugal, and Spain. Their absence was noted very early as a major shortcoming by scholars claiming a distinct Latin-rim or Mediterranean model (Ferrera 1996; Leibfried 1992).

Ironically when *Three Worlds* was published in 1990, the fall of communism had opened up the previously socialist Central and Eastern Europe (CEE) to liberal democratization and free market capitalism. Recent research seeks to extend the regime typology to the New Europe, studying the transition of post-socialist social protection as they came under the influence of international organizations and as a dozen countries joined the European Union in 2004 and 2007. Nevertheless quantitative analyses of advanced welfare states rarely include these CEE countries due to their still-unsettled situation as transition countries (for an exception see Castles and Obinger 2008).

Also with regard to the Anglophone nations, Esping-Andersen's typology has been immediately criticized. Castles and Mitchell (1993: 105) distinguish a fourth "radical" welfare regime that combines low social expenditure with achieving

Table 1: Esping-Andersen's Welfare Regimes and Further Comparative Studies

		All 13 typologies				2 typologies by Esping-Andersen		Other 11 typologies (*)			
		Coverage		Consistency				Coverage		Consistency	
		N	(%)	N	(%)	(1990)	(1999)	N	(%)	N	(%)
Continental	Germany	13	100%	10	77%	<u>Con</u>	<u>Con</u>	11	100%	8	73%
	Austria	12	92%	9.5	79%	<u>SD/Con</u>	<u>Con</u>	10	91%	8	80%
	France	12	92%	12	100%	<u>Con</u>	<u>Con</u>	10	91%	10	100%
	Belgium	11	85%	6.5	59%	<u>SD/Con</u>		10	91%	6	60%
	Italy	11	85%	8	73%	<u>Con</u>	<u>Con</u>	9	82%	6	67%
	Spain	5	38%	3	60%		<u>Con</u>	4	36%	2	50%
	Portugal	3	23%	2	67%			3	27%	2	67%
	Greece	3	23%	1	33%			3	27%	1	33%
	Czech Republic	1	8%	1	100%			1	9%	1	100%
Liberal	Australia	13	100%	9	69%	<u>Lib</u>	<u>Lib</u>	11	100%	7	64%
	Canada	13	100%	12	92%	<u>Lib</u>	<u>Lib</u>	11	100%	10	91%
	United Kingdom	13	100%	7.5	58%	<u>Lib</u>	<u>(Lib)</u>	11	100%	6	55%
	United States	13	100%	13	100%	<u>Lib</u>	<u>Lib</u>	11	100%	11	100%
	New Zealand	11	85%	6	55%	<u>Lib</u>	<u>Lib</u>	9	82%	4	44%
	Ireland	10	77%	4.5	45%	<u>Lib/Con</u>		9	82%	4	44%
	Japan	10	77%	5.5	55%	<u>Con/Lib</u>	<u>Con</u>	8	73%	5	63%
	Switzerland	8	62%	4.5	56%	<u>Con/Lib</u>		7	64%	4	57%
	Iceland	1	8%	1	100%			1	9%	1	100%
Social-democratic	Denmark	13	100%	13	100%	<u>SD</u>	<u>SD</u>	11	100%	11	100%
	Norway	13	100%	13	100%	<u>SD</u>	<u>SD</u>	11	100%	11	100%
	Sweden	13	100%	13	100%	<u>SD</u>	<u>SD</u>	11	100%	11	100%
	Finland	12	92%	6.5	54%	<u>Con/SD</u>	<u>SD</u>	10	91%	5	50%
	Netherlands	12	92%	5	42%	<u>SD</u>	<u>SD</u>	10	91%	3	30%
Central and Eastern Europe	Estonia	1	8%	1	100%			1	9%	1	100%
	Hungary	1	8%	1	100%			1	9%	1	100%
	Latvia	1	8%	1	100%			1	9%	1	100%
	Lithuania	1	8%	1	100%			1	9%	1	100%
	Poland	1	8%	1	100%			1	9%	1	100%
	Slovenia	1	8%	1	100%			1	9%	1	100%
	Slovak Republic	1	8%	1	100%			1	9%	1	100%

Notes: Categorization—Lib: Liberal (or residual); Con: Conservative (or corporate, European, South); SD: Social Democratic (or universalist); not shown: “radical” and “undefined”; Esping-Andersen (1990): decommodification/stratification (split categorization: 0.5); dominant regime (consistency index) is underlined.

Source: Own calculations and coding based on Esping-Andersen (1990, 1999) and (*) 11 studies cited in Arts and Gelissen (2010); CEE countries (incl. Czech Republic) covered only in Castles and Obinger (2008).

equality through labor relations, including Australia and New Zealand (but also the United Kingdom before 1980). In addition, the disregard for Japan's singularity and the lack of discussion of the rapidly industrializing East-Asian economies was thought to be an omission of the threefold typology. Esping-Andersen (1999) revisited his typology a decade later (see Table 1), including Spain but dropping hybrid cases such as Belgium, Ireland, and Switzerland. Thus the 1999 revision in fact reduced the number of cases from 18 to 16 OECD countries, still excluding the new EU-accession countries and Asian “tiger” economies.

Comparative studies that tested Esping-Andersen's typology often used a similar set of OECD countries (see Table 1). The example of Esping-Andersen's regime typology and its

reaction by area specialists indicates that the number of ideal-type models of welfare states is contested. Esping-Andersen has been defending his threefold typology and claims that adding further types will dilute the analytical power at the risk of having as many types as cases (Esping-Andersen 1993). Others scholars have pointed at the exceptionalism of particular cases or extended the initial set of countries, thereby adding other European or world regions.

Several comparative analyses have been undertaken to test the clustering along the Three World typology (Arts and Gelissen 2002, 2010). Different methods and data sources have been applied, ranging from descriptive face-value tabular analysis to formal Qualitative Comparative Analysis to quantitative cluster or principal component analysis. The meta-analysis in

Table 1 reports the coverage of countries and their categorization by Esping-Andersen (1990, 1999) and further comparative studies that were published since the 1990s (Arts and Gelissen 2010). In these 11 follow-up studies the number and set of countries varies considerably from a low of 11 countries to as many as 25 countries, though most range around Esping-Andersen's 16 to 18 OECD countries.

Some countries were always included, such as Britain, Germany, and Sweden, while others might be included rather rarely, in particular the Southern European (only three to five times) and CEE countries (only once), not to speak of some new OECD countries outside Europe (e.g., South Korea and Mexico) that were commonly ignored. Among those countries with more than eight studies (ten including Esping-Andersen's), the consistency of categorization is perfect or relatively high in the case of some prime cases: in particular Sweden (also Denmark and Norway) as a Social Democratic welfare state and the USA as the Liberal model, while three studies diverted from placing Germany as the prime case of the Conservative regime. There are some cases in which there is not much consistency, such as Finland and the Netherlands; these are very hybrid mixes of several regime types. Thus there are some welfare states that stand out in fitting the ideal-type most closely, while many others are classified more ambiguously depending on the indicators used, time frame, and selection of cases. The selection of countries indeed matters.

Country Selection in Quantitative Comparisons

The example of Esping-Andersen's *Three Worlds* and its follow-up studies show that comparative welfare state research tends to rely on medium-N research designs. This involves some half-dozen to two dozen countries that are by and large hand-selected from the members of international organizations, in particular the European Union and/or the OECD. As the case of welfare regimes indicates, it is not only the number of cases that matters but also the particular criteria for including or excluding these. The selection of cases in comparative research is a major decision with important consequences not only for the external validity but also internal validity of macro-comparative analysis (Ebbinghaus 2005; Geddes 1990).

There are theoretical, substantial and practical methodological reasons to select member countries of international organizations. First of all, the OECD and EU are seen as clubs of members that share enough commonality to make comparison possible, for instance, countries with similar economic development (OECD) or that have substantial political integration (EU). By December 2010, the OECD had accepted 34 countries (13 outside the EU), while the European Union has had 27 member states since 2007 (of which six are not yet OECD members). Note that both international organizations have substantially increased membership since the 1990s, thus cross-national time-series analyses before 1990 cover often fewer cases than those of more recent years. The current 40 combined cases of OECD and/or EU member countries represent more than 20% of the 192 UN member-states, and a much larger percentage in economic power and political weight. It is obvious that the OECD's current 34 members might be a large

medium-N, yet we can hardly generalize findings from these advanced open economies to the less-developed rest of the world. We need to acknowledge these scope conditions for any generalization of welfare regime analysis.

A major problem of medium-N comparison is thus the selection bias posed by the historically and politically given cases for quantitative analysis (Ebbinghaus 2005). "Selection bias is commonly understood as occurring when the non-random selection of cases results in inferences, based on the resulting sample, that are not statistically representative of the population" (Collier 1995: 462). While it is doubtful whether we can speak of a random sample in cross-national comparison, and thus inference statistics should be used, descriptive statistics can be applied to a given set of cases.

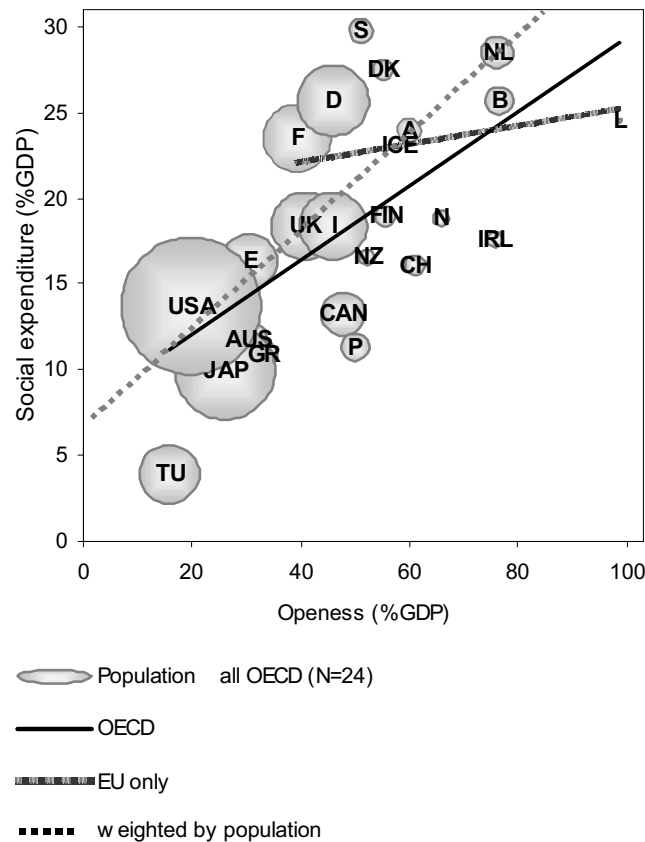
As Ragin (2000: 46) observes, "having a reasonably well-delimited population is a precondition for the quantitative analysis of cross-case patterns." Yet in practice, variable-oriented researchers choose a set of cases for largely pragmatic reasons: cultural familiarity, similar socio-economic development and—last but not least—availability of data. Cross-national studies are commonly dealing with a deliberate selection from a categorical set of political units. Moreover, data availability is unequally distributed across potential cases: macro-indicators and time-series are more likely to be available for the more advanced economies, larger societies, and long-term democracies, in particular as the OECD or Eurostat collect such data-series.

In contrast to a population census, the cases of a macro-comparison are heterogeneous social units, violating the homogeneity assumption of inferential statistics. Even when we compare the OECD before the recent opening up to new members, such a selection of 24 countries is highly stratified: The member-states range from tiny Iceland to the 1,000-times-larger USA, from rich Switzerland to four times poorer Turkey (in terms of GDP per head). Given these large differences in population and economic resources, it may be misleading to analyze each case as equally important. Indeed it would be better to use arithmetic averages weighted by population in many of our comparisons.

For illustration, Figure 1 plots the relationship between social expenditure and openness of an economy, testing Katzenstein's thesis (1985) of small states' effects on welfare state expansion. The arithmetic average of social expenditure (in % GDP) for the 24 OECD countries in 1980 is 18.6%, varying between Turkey (4%) and Sweden (30%), whereas the (population) weighted average is 15.8% due to the low spending of the larger countries (particularly the USA with only 14% of GDP, yet it represents 30% of OECD population in 1980). The diversion is even larger in relation to economic openness (average of imports and exports in % GDP) since the larger the country, the bigger the domestic market. The most common fix to this problem is to apply statistical controls for measuring (otherwise unobserved) heterogeneity, including population or resource variables as controls in a regression model. However, adding control variables will considerably reduce the degrees of freedom in cases where cross-country analysis are based on medium-N designs and might limit the explanatory

Figure 1: Social Expenditures and Openness in OECD Countries (1980s)

Source: Ebbinghaus (2005)



power of other independent variables in case of multicollinearity.

The alternative strategy sometimes undertaken is to reduce heterogeneity by selecting countries with similar conditions (e.g., advanced industrial countries) or by agnostically assuming similarity due to membership in a particular dataset (e.g., OECD). Both strategies have major repercussions in descriptive and inference statistics. For instance, by using the EU member-states as of 1995 instead of OECD's 24 of 1990, the regression line in Figure 1 is much flatter, indicating that the Katzenstein thesis holds less within Europe than across the OECD. Hence, there are two problems that cross-sectional quantitative analysis faces: The population of macro-units is very heterogeneous, contradicting the homogeneity assumption, and any selection thereof, be it for size or categorical membership, will have considerable impact on dependent and independent variables (selection bias). These problems result from the historical contingency of real-world macro-social units.

When selection is based on a regional (e.g., Southern Europe), categorical (e.g., Social Democratic welfare states) or organizational membership (e.g., EU), the full set of real-world cases is biased as a result of historical and political contingencies. Members of a political organization (e.g., EU) have to fulfill the admission criteria, while the accession is ultimately a political decision. Even when we use regional or categorical sets or for that matter all existing countries, the universe of

cases remains biased. All observable cases represent only a "limited diversity" (Ragin 1987), that is, we can observe less diversity than is logically possible for the variables under investigation. This has not only repercussions for testing potential hypotheses with the limited set of real-world cases, but this contingency also reduces the inference from frequency-related statistics.

A counterfactual example, throwing the dice of national state building, may suffice to show the impact of past history on today's available cases. What would happen to OECD regression analysis of welfare state expenditures if the Nordic countries had been all united, instead of remaining independent cases of high welfare spenders? Instead of five out of 24 OECD countries, there would be only one Nordic Union out of then 20 OECD data points and the regression line would be much flatter for any indicator of social democratic power. When we assume universal causality and test this with cross-sectional linear regression, the historical contingency might indeed affect systematically the frequency of confirming and disconfirming cases. This is the contingency bias of past nation-state building that affects our observable real-world cases.

Typologies and Qualitative Comparative Analysis

Are configurational methods a solution for the shortcomings of comparative research on welfare states? Since the cases available for comparative analysis—the advanced welfare

states—are themselves the result of macro-historical factors, quantitative (general linear) regressions are less suited to deal with this selection problem than configurational analysis. In Qualitative Comparative Analysis, the macro-configurations (for instance the type of national states, the religious tradition and degree of international integration) can be taken into account in order to explain a particular outcome (here a particular type of welfare state regime). For instance, the question whether all Social Democratic welfare states are protestant nation states is specific to a subset of cases under investigation, it seems a necessary but not sufficient condition. There is also the possibility to allow for *multifinality*; for example, protestant societies could be either Liberal or Social Democratic welfare states depending on further intervening conditions.

In an empirical test of Esping-Andersen's welfare regime, Ragin (1994) used first a cluster analysis to code the dependent variable (regime type) based on Esping-Andersen's (pension system) indicators and the (multi-value) QCA method to explain the different regime types (three dummy variables for the regimes but also allowing for an undetermined outcome in 6 out of 18 countries). Thus the different regime configurations can be modeled by dummies for each regime type, based on qualitative coding and/or statistical methods. As long as the number of cases is larger than the number of types, QCA could be used to relate configurational conditions to the outcome variable. There is no predetermined number of cases, but we would need to legitimate distinguishing an additional type. Thus (multi-value) QCA should not be used as an inductive method to derive a typology but to sort out configurational factors associated with a particular outcome (an ideal-type regime).

A major issue is whether typologies are conceptualized as exclusive categories (that is, either/or conditions) or can be thought of as partial membership, in particular for those hybrid regime cases. In the former case (crisp-set or multi-value) QCA is appropriate if there is no problem to code qualitative or reduce quantitative information to binary coding; in the latter case (and if measurement problems occur) it is advisable to use fuzzy set analysis (Ragin 2000; Rihoux and Ragin 2009) since it allows modeling of partial membership in non-exclusive typologies. Thus we can measure Liberal, Social Democratic and Conservative characteristics as independent dimensions. A welfare state could be partially member of all three categories: They could be located anywhere within a triangle between the three ideal-type extremes. For instance, such a fuzzy set analysis of reform processes in Nordic welfare states indicates that there are notable differences across rather similar countries, between social policy areas and over time (Kvist 1999). Today's Nordic welfare systems only partially concur with the ideal-type Social Democratic model of Esping-Andersen. Fuzzy set analysis has thus far only rarely been applied to comparative welfare state research (Emmenegger 2011; Pennings 2005; Vis 2009).

Fuzzy set analysis may be particularly appropriate for the comparative analysis of medium-N cases, such as the advanced welfare states. Fuzzy sets can be used to better capture the partial membership in theory-driven typologies that are then

measured by qualitative and quantitative indicators. Such fuzzy set typologies of regimes can be applied at different levels, as overall features of the welfare regime but also as particular features of subsystems such as pension or employment policy. While the overall typology of welfare regimes has been crucial in providing some overall summary measures of welfare states' main tendency, more fine-grained analyses of particular policy areas have showed the need for more adapted typologies to capture policy-specific features. This has also shed light on why some welfare states seem to be more hybrid than others, as they combined different types of welfare policies in the specific policy areas. For example, Italy and France share similar Bismarckian pension systems, but in health care France has a Bismarckian sickness scheme, while Italy a Beveridge-type national health system.

While comparative welfare state analysis has used regime typologies as the explanandum, it can also serve as the explanans for particular outcome variables such as inequality. As an independent variable, the common welfare regime typology was often used as one of many macro-variables in quantitative models to explain the dependent variable of interest. Again, fsQCA could serve as a tool to select the appropriate independent variables. Given that in time-series analyses or in multi-level cross-national analysis (with individual level data) only few macro-variables can be tested, it is important to have well-established typologies of welfare state regimes that can be used by non-experts as one of competing macro-variables. In addition, it is advisable to use more fine-grained typologies of particular policy areas to model more closely the underlying causal mechanisms. For instance, in order to explain differences in inequality measures of the redistributive nature of welfare states (such as the decommodification index of Esping-Andersen) maybe more relevant than a summary regime classification.

Selection Matters in Comparative Analysis

In this essay, welfare state regime analysis à la Esping-Andersen was taken as an example to show how comparative analysis faces some important methodological challenges when typologies are developed on a selection of country cases. The analysis of advanced welfare states is limited by a pre-given set of observable cases that are the result of historical and political processes. Thus the case selection is often not independent of the causal factors explaining the differences across welfare states. Since case selection in cross-national analyses is all but a random sample from the population at risk, comparative analysis faces a particular selection bias. We need to be more explicit about the rationale for selecting cases and specify the scope conditions when making generalizations beyond the observed cases.

Esping-Andersen's initial regime typology was based on 18 OECD countries, though its main insights were drawn from some prime examples (Germany, Sweden, Britain, and USA). The threefold typology has been criticized for not paying attention to finer differences and by not taking into account the Southern European, Eastern European, and Asian peculiarities of welfare state development. The meta-analysis of compara-

tive studies on welfare regimes (see Table 1) indicates that the number and coverage of cases varies considerably between studies, not to mention the differences in the applied methods and indicators used. Thus not all countries fit neatly into the ideal-type categories, and some cases have often been ignored.

More generally, the essay discussed the problems of generalizing from a selection of advanced welfare states (OECD or EU) in comparative welfare state research. The very process of welfare state development is closely related to the macro-historical and political processes leading to economic growth and political integration that are preconditions for membership in these international organizations. Therefore comparative welfare state analysis faces a selection problem: The causal factors explaining welfare state development are also related to the factors determining the available cases for analysis. Therefore configurational comparative methods seem better suited than cross-national quantitative methods such as linear regression analysis. Configuration analysis can be used to analyze subsets of causal factors explaining specific typologies. Although multi-value QCA can be used to code welfare state typologies based on qualitative and quantitative indicators, it is advisable to use fuzzy set QCA to allow for a more open conceptualization of welfare regime typologies. Typologies can be used then as membership set relations not unlike the conceptualization of a comparison of ideal-type and real-type observations as advocated by Max Weber.

References

- Arts, Wil and John Gelissen. 2002. "Three Worlds of Welfare Capitalism or More? A State of the Art Report." *Journal of European Social Policy* 12:2, 137–158.
- Arts, Wil and John Gelissen. 2010. "Models of the Welfare State." In *The Oxford Handbook of the Welfare State*. Francis G. Castles, Stephan Leibfried, Jane Lewis, Herbert Obinger, and Christopher Pierson, eds. (Oxford: Oxford University Press), 569–583.
- Castles, Francis G. and Deborah Mitchell. 1993. "Worlds of Welfare and Families of Nations." In *Families of Nations: Patterns of Public Policy in Western Democracies*. Francis G. Castles, ed. (Aldershot: Dartmouth), 93–128.
- Castles, Francis G. and Herbert Obinger. 2008. "Worlds, Families, Regimes: Country Clusters in European and OECD Area Public Policy." *West European Politics* 31:1/2, 321–344.
- Collier, David. 1995. "Translating Quantitative Methods for Qualitative Researchers: The Case of Selection Bias." *American Political Science Review* 89:2, 461–466.
- Ebbinghaus, Bernhard. 2005. "When Less is More: Selection Problems in Large-N and Small-N Cross-National Comparison." *International Sociology* 20:2, 133–152.
- Ebbinghaus, Bernhard and Philip Manow, eds. 2001. *Comparing Welfare Capitalism: Social Policy and Political Economy in Europe, Japan and the USA*. London: Routledge.
- Emmenegger, Patrick. 2011. "Job Security Regulations in Western Democracies: A Fuzzy Set Analysis." *European Journal of Political Research* 50:3, 336–364.
- Esping-Andersen, Gøsta. 1990. *Three Worlds of Welfare Capitalism*. Princeton, NJ: Princeton University Press.
- Esping-Andersen, Gøsta. 1993. "The Comparative Macro-Sociology of Welfare States." In *Social Exchange and Welfare Development*. Luis Moreno, ed. (Madrid: Csic), 123–136.
- Esping-Andersen, Gøsta. 1999. *Social Foundations of Postindustrial Economies*. Oxford: Oxford University Press.
- Ferrera, Maurizio. 1996. "The 'Southern Model' of Welfare in Social Europe." *Journal of European Social Policy* 6:1, 17–37.
- Geddes, Barbara. 1990. "How the Cases You Choose Affect the Answers You Get: Selection Bias in Comparative Politics." In *Political Analysis*, James A. Stimson, ed. (Ann Arbor: University of Michigan), 131–150.
- Hall, Peter A. and David Soskice, eds. 2001. *Varieties of Capitalism: The Institutional Foundations of Comparative Advantage*. New York: Oxford University Press.
- Katzenstein, Peter J. 1985. *Small States in World Markets. Industrial Policy in Europe*. Ithaca, NY: Cornell University.
- Kvist, Jon. 1999. "Welfare Reform in the Nordic Countries in the 1990s: Using Fuzzy-Set Theory to Assess Conformity to Ideal Types." *Journal of European Social Policy* 9:3, 231–252.
- Lange, Peter and Hudson Meadwell. 1991. "Typologies of Democratic Systems: From Political Inputs to Political Economy." In *New Directions in Comparative Politics*. Howard J. Wiarda, ed. (Boulder, CO: Westview), 82–117.
- Leibfried, Stephan. 1992. "Towards a European Welfare State? On Integrating Poverty Regimes into the European Community." In *Social Policy in a Changing Europe*. Zsuzsa Ferge and Jon Eivind Kolberg, eds. (Frankfurt: Campus), 245–279.
- Pennings, Paul. 2005. "The Diversity and Causality of Welfare State Reforms Explored with Fuzzy Sets." *Quality & Quantity* 39, 317–339.
- Ragin, Charles C. 1987. *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*. Berkeley: University of California Press.
- Ragin, Charles C. 1994. "A Qualitative Comparative Analysis of Pension Systems." In *The Comparative Political Economy of the Welfare State*. Thomas Janoski and Alexander M. Hicks, eds. (New York: Cambridge University Press), 320–345.
- Ragin, Charles C. 2000. *Fuzzy-Set Social Science*. Chicago: University of Chicago Press.
- Rihoux, Benoît and Charles C. Ragin, eds. 2009. *Configurational Comparative Analysis. Qualitative Comparative Analysis (QCA) and Related Techniques*. Los Angeles: Sage.
- Vis, Barbara. 2009. "The Importance of Socio-economic and Political Losses and Gains in Welfare State Reform." *Journal of European Social Policy* 19:5, 395–407.

One-Size-Fits-All? Measurement Issues in Medium-N Comparative Welfare State Analysis

Jon Kvist

University of Southern Denmark
kvi@sam.sdu.dk

Medium-N studies often aim to explore diversity, i.e., the simultaneous existence of similarities and differences across cases (Ragin 2008). Diversity-oriented analysis is vividly illustrated in comparative welfare state analysis, where the typology of welfare states advocates a configurational view of cases (Esping-Andersen 1990). For example welfare states are differentiated according to whether they are universal or selective on one dimension and generous or ungenerous on another dimension.

However, when conducting medium-N studies welfare state scholars often resort to measures that are not true to the meaning of configurations or to the theoretical concepts they intend to measure, and in some studies measures that are not comparable across cases or over time. Scholars appear to have too many cases to make informed cross-case comparisons and too few to make robust statistical analyses. The question thus arises as to whether medium-N studies sit uncomfortably between small-N and large-N studies when it comes to measurement issues.

What do scholars do when they have a total population of between 18 and 35 countries that meaningfully qualify as welfare states? Most scholars use one of two dominant approaches. One approach is for scholars to resort to comparing only a few countries, perhaps even single-case studies, and relating their empirical findings to theory building and other empirical findings in comparative welfare state research. As scholars using this approach are no longer in the field of medium-N studies, I will not consider their particular measurement challenges here.

Instead I will focus on the second approach, that of those who believe in “one size fits all.” This approach is namely to use the measures of large-N studies in medium-N comparative welfare state analyses. Typical large-N measures include aggregate measures like social expenditures, single-case indicators such as measures for a specific stylized type, averages for different groups of countries, and composite indexes covering different aspects. In such studies scholars use large-N measures and data as though these have been designed to accommodate a wide range of questions and research designs.

I argue that the practice of using large-N measures in medium-N studies is problematic with regard to the dual challenge of securing content validation and ensuring comparability (for more on these issues of measurement validity, see Adcock and Collier 2001). Securing content validation entails understanding how we can adequately capture the meaning of theoretical concepts and analytical constructs. Ensuring comparability entails making measures that are comparable with measures in other contexts. To illustrate the problems of using aggregate large-N measures in medium-N studies, I offer two examples. Each example focuses on a specific way of theoretically operationalizing the welfare state, namely welfare effort and benefit entitlements. Each example uses different measures of welfare states typically used in large-N studies. Together the two examples cover aggregate measures, single-case indicators, group averages and composite indexes that all have specific problems in meeting the challenges of content validation and comparability. But first I will sketch how different types of welfare states are conceptualized and measured.

Measuring Different Types of Welfare States

Comparative welfare state research is rich on theories and concepts. The conventional starting point is the three worlds of welfare capitalism as advanced by Gøsta Esping-Andersen (1990, 2009). Known as the Liberal, Conservative-Corporatist, and the Social Democratic types, these welfare state regimes are constituted by different factors that scholars have opera-

tionalized for inputs and outputs.

The dominant input aspect is welfare effort as measured by social expenditures as percentage share of gross domestic product (GDP). The dominant output factor is social rights, or benefit entitlements, which derive from the ways that the institutional design of policies allows access to benefits and stipulate benefit generosity.

In short, countries with a Liberal welfare regime provide welfare benefits only as a last resort, when the market and families have failed. We expect low social expenditures, with benefits that are selective and not generous.

Countries with the Conservative-Corporatist welfare regime provide good coverage to the labor market insiders and meager benefits to outsiders. We expect medium social expenditure levels, as well as generous benefits for insiders and ungenerous benefits for outsiders.

Countries with a Social Democratic welfare regime have universal coverage and generous benefits. We expect high social expenditures, with benefits that are particularly generous for low-income groups, generous for middle-income groups, and the least generous for high-income groups.

How do scholars capture these constitutive factors in their measurements? And what issues related to content validation and comparability do these measurements raise?

Social Expenditures

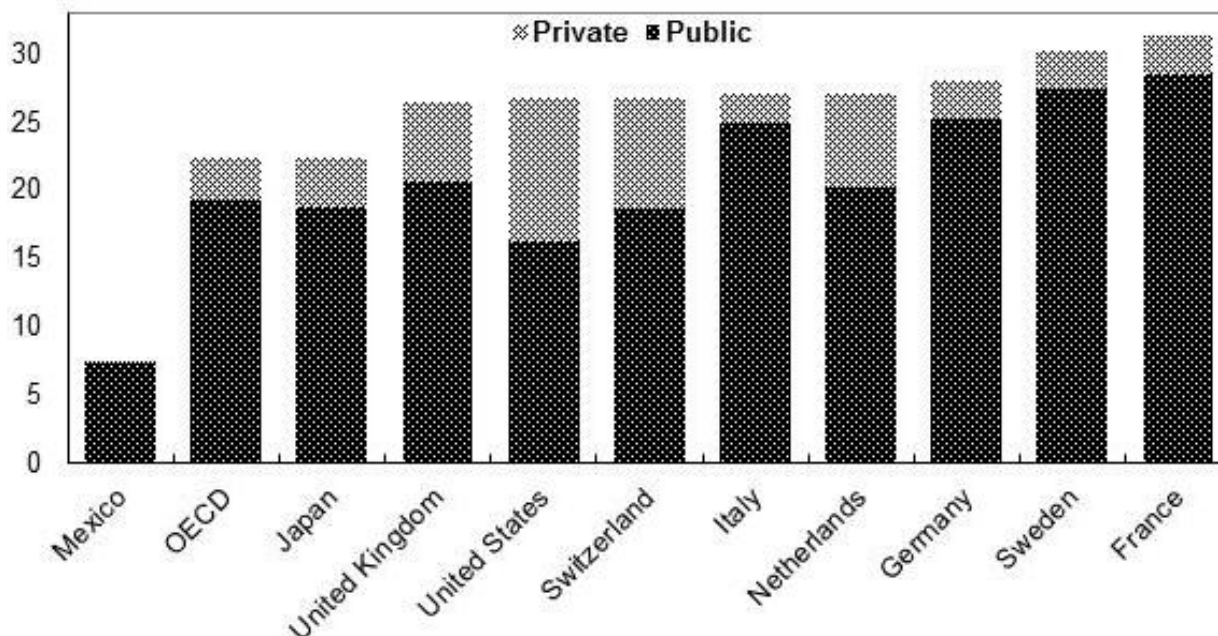
Public social expenditure as a percentage of GDP is by far the most frequent measure in comparative welfare state research. Arguably, the measure has proven effective in large-N studies for distinguishing between countries with a welfare state and those without a welfare state. Figure 1 shows the difference in social expenditure in ten countries and the average for countries that are members of the international Organisation for Economic Co-operation and Development (OECD). With 7.4 percent of GDP devoted to social expenditures, Mexico does not qualify for the group of countries with welfare states. Social expenditures in the countries with welfare states range between 22.0 percentage of GDP as in Japan (and the OECD average) and 31.3 percentage of GDP as in France.

However, social expenditures do not allow for accurate distinctions between countries with different types of welfare states, nor does social expenditure reflect the meaning of welfare effort as understood by researchers or people in general.

Public social expenditure cannot differentiate between countries with different types of welfare states. Figure 1 shows that Sweden with a Social Democratic welfare state type has high public social expenditures as expected. But at the same level we also find France, Germany, and Italy, which are associated with the Conservative-Corporate welfare state type. Sweden and Italy appear the same even though any scholar even slightly familiar with social policy in the two countries knows that they are very different: The Italian welfare state has lavish old-age pensions for labour-market insiders, but rudimentary benefits in almost all other areas. The Swedish welfare state has extensive, universal benefits across a wide range of areas. Social expenditures are not comparable across cases; a per-

Figure 1: Public and Private Social Expenditure as Percentage of GDP, 2007

Source: OECD (2011)



centage point of GDP is seen as the same in the two countries although their redistribution effects and allocation of expenditures on policy areas and programs differ markedly.

Public social expenditure does not encompass all relevant aspects of welfare state expenditures. As the OECD points out, social expenditures must take into account different tax systems and private social expenditures (Adema and Ladaïque 2009). Switching from gross public social expenditure to net public and private social expenditure has two important implications. First, social expenditure levels rise, sometimes dramatically, as in the U.S. Second, cross-national differences diminish. As Figure 1 shows, the Netherlands and Switzerland now have the same expenditure level as countries normally associated with the Conservative-Corporative welfare state-type—and so does the U.S.

Another weakness of social expenditure as a measure of welfare effort is that it aligns neither with scholarly interest in the politics of welfare reforms nor with people’s perceptions of well-being. As Esping-Andersen (1990) notes, that actors have ever pressed for more social expenditure per se is hard to believe. When interpreted as an indicator of welfare effort, Finland had the biggest welfare state in the mid-1990s but less at the start and end of the decade, an observation that squared badly with public perception at the time. In the 1990s Finland hit an economic crisis in the first half of the decade, recovering in the second half. Public social expenditure varied from 24.1 percent of GDP in 1990, over 30.7 in 1995 to 24.2 in 2000 (OECD 2011). The measure depends as much on the development in the economy, GDP, as in social expenditure. Social expenditures do not pass the test of content validation.

In sum, while social expenditure may have an explanatory power when used in large-N studies, as soon as we limit the analysis to a medium number of countries, the explanatory

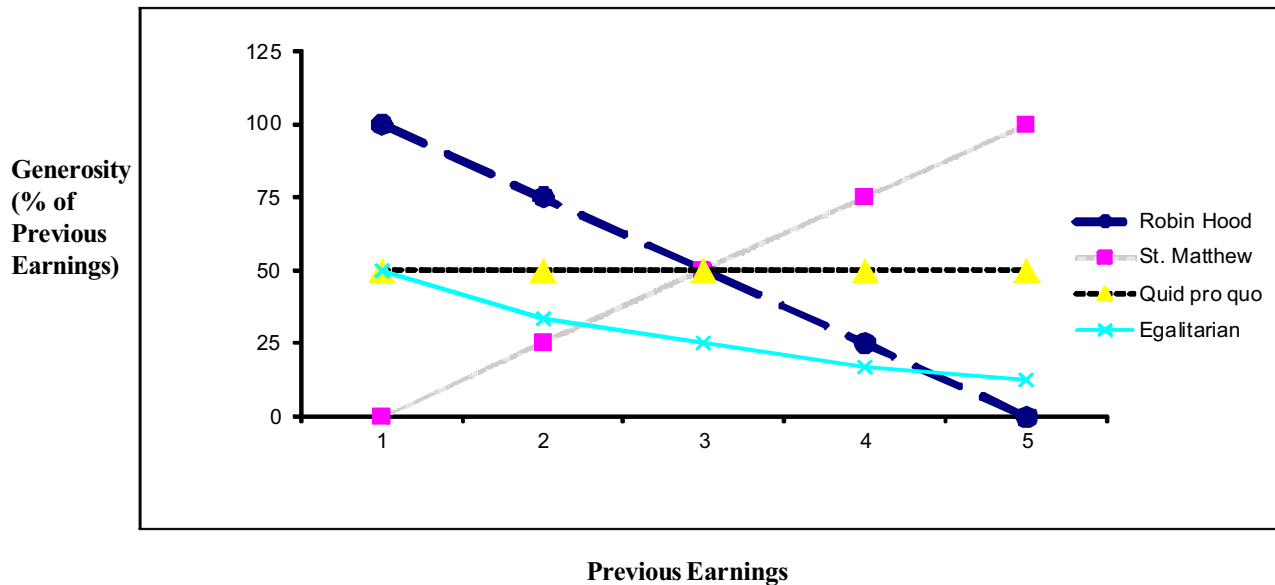
power diminishes. To establish comparability, measures must take account of different tax treatment and provision of benefits in both the public and the private sector. After accounting for private expenditures and the tax system, we see how countries with welfare states range within a few percentage points, making it difficult to make clear distinctions between countries. Moreover, expenditure levels do not reflect the concern in many comparative welfare state analyses such as tapping into the total welfare effort or the redistributive strategies of who benefits. Finally, expenditure measures do not unmask differences in the allocation of means across different policy areas and programs. Scholars using expenditures as a measure for welfare state efforts must take account of all these concerns if their analyses are to yield meaningful results.

Benefit Entitlements

As described earlier, there are distinct theories about how the generosity of benefits varies for different socio-economic groups in different types of welfare states. I will illustrate the heterogeneous impact of the welfare state by referring to the expected profiles of benefit generosity according to four strategies of redistribution, as set out in Figure 2. The Robin Hood strategy is to take from the rich and give to the poor, resulting in a downward slope with benefits being very generous to people with low previous earnings and ungenerous benefits for high-income groups. The complete opposite is the case in the St. Matthew strategy, where more is given to those who have. The quid pro quo—or “something for something”—strategy implies that one receives according to what one pays in. Finally, the egalitarian strategy gives something to everybody, which translates into more generous benefits for the poor than for the rich.

We can expect that European countries with a Liberal wel-

Figure 2: Four Strategies of Redistribution Measured by the Generosity of Social Benefits as a Percentage of Previous Earnings



fare state type have a profile following the egalitarian strategy (while this may not be the case for non-European countries with a Liberal welfare state type, we need not enter this debate now). Countries with a Conservative-Corporative welfare state type focus on status maintenance, thus following the quid pro quo strategy. Countries with a Social Democratic welfare state type that both is heavy on redistribution and gives generous benefits to middle-income groups have a profile that follows the Robin Hood strategy for low incomes and the quid pro quo strategy for higher incomes.

Single-Case Indicators

To measure benefit entitlements, scholars making medium-N studies use single-case indicators, composite indicators, or group indicators as large-N studies often do. The situation of the average production worker is the single-case indicator that is used in the two perhaps most important databases, i.e., the Social Citizenship Indicator Program (SCIP) and the Welfare State Entitlement Project (WSEP) (Korpi and Palme 2007, Scruggs 2004). As the average production worker becomes extinct, the comparability of measures over time becomes difficult. However, there is also a more general problem potentially associated with single-case indicators, namely that it is not possible to generalize from the situation of the single-case to other cases. To illustrate this, I have calculated net replacement rates for persons with previous earnings varying between 75% and 200% of the average production worker.

Figure 3 shows that cross-national differences increase with previous earnings. The gap between the countries with the lowest and highest net replacement rate increases from 8 to 39 percentage points from the lowest previous earnings level to the highest. At the level of the average production worker, the gap is 17 percentage points. Single-case indicators of social citizenship provide a partial picture of cross-

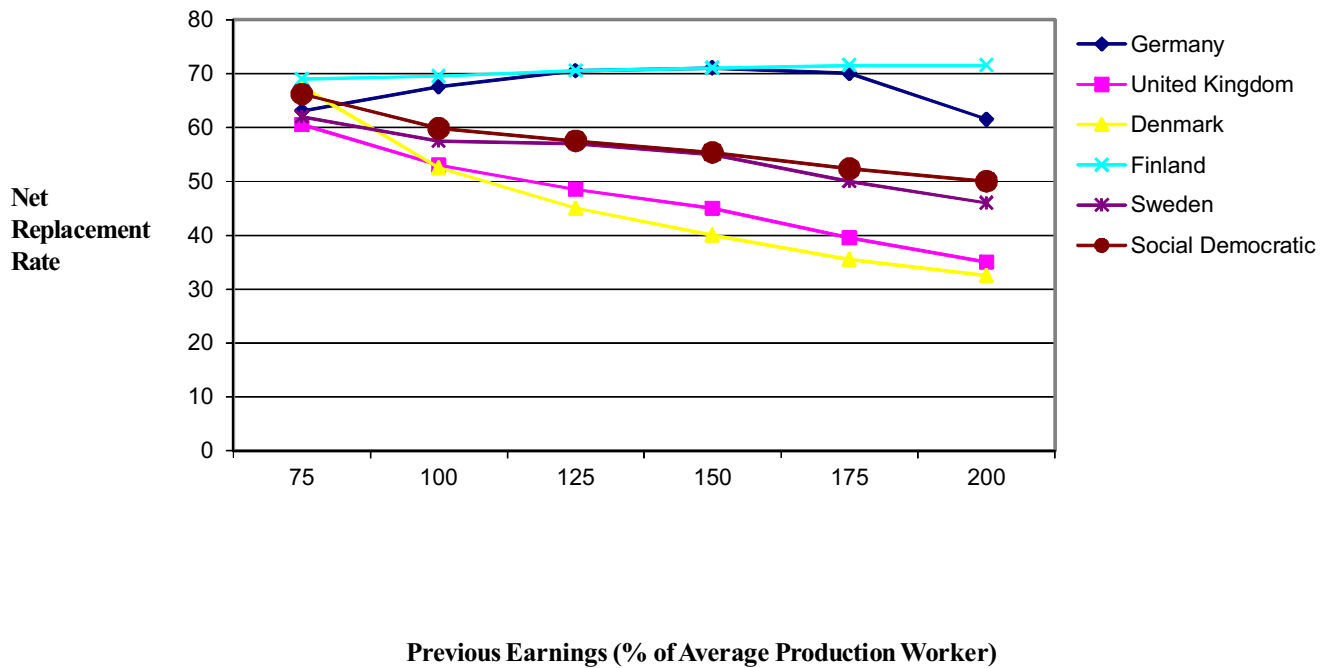
national differences.

Nevertheless, single-case indicators may not provide the most relevant basis for advancing theory or advocacy of policy. If poverty among the elderly is the issue of concern, then the situation of persons with low previous earnings may be more relevant than that of the average production worker. If sustainability of pension systems is the issue of concern, then the situation of persons with high previous earnings may be the most appropriate to examine. For example, Danish full old-age pensions are the second highest for low-income groups, but the lowest for high-income groups. The Danish plan thereby looks effective in reducing poverty in old age as well as sustainable in economic terms. However, the measure only covers public benefits. Just as we saw that private social expenditures matter for the ranking of countries, such private programs also matter for benefit entitlements. In the Danish example, including private pensions would raise the level of benefits for middle- and high-income groups and modify the picture of a sustainable pension system.

The Finnish scheme comes close to the quid pro quo strategy. Because pension income is not taxed or is taxed more lightly than income from work, the German profile shows elements of the St. Matthew strategy, in that replacement rates increase until the plan is capped at an income level approximately 175% of that of the average production worker.

Making group averages masks intra-group differences. Comparative studies of social insurance tend to claim that countries belonging to the Conservative regime have the highest generosity and countries belonging to the Liberal regime the least generous. The Social Democratic countries are often said to occupy a middle position (see, e.g., Scruggs 2006). Figure 3 shows how group averages are problematic. The profile for the Social Democratic regime made up by the average for the three Nordic countries in the study conforms nicely to the expecta-

Figure 3: Full Public Old-Age Pensions in Five European Countries as a Percentage of Previous Earnings, Single Person, 2006



tion across all income levels. However, in reality the Nordic countries are all over the map, with Finland being more generous than Germany, and Denmark being less generous than the UK for middle- and high-income earners.

Composite Indicators

Composite indicators come in two forms. Perhaps the most common composite indicator is an average of scores across two or more typical cases. Both databases on social rights and welfare entitlements (SCIP and WSEP) give the average net replacement rate for the standard and the minimum old age pensions. The other type of composite indicator is an indicator based on a variety of different types of information, typically put together in an index. Examples include the de-commodification indicator of Esping-Andersen (1990) and the generosity indicator of Scruggs (2006).

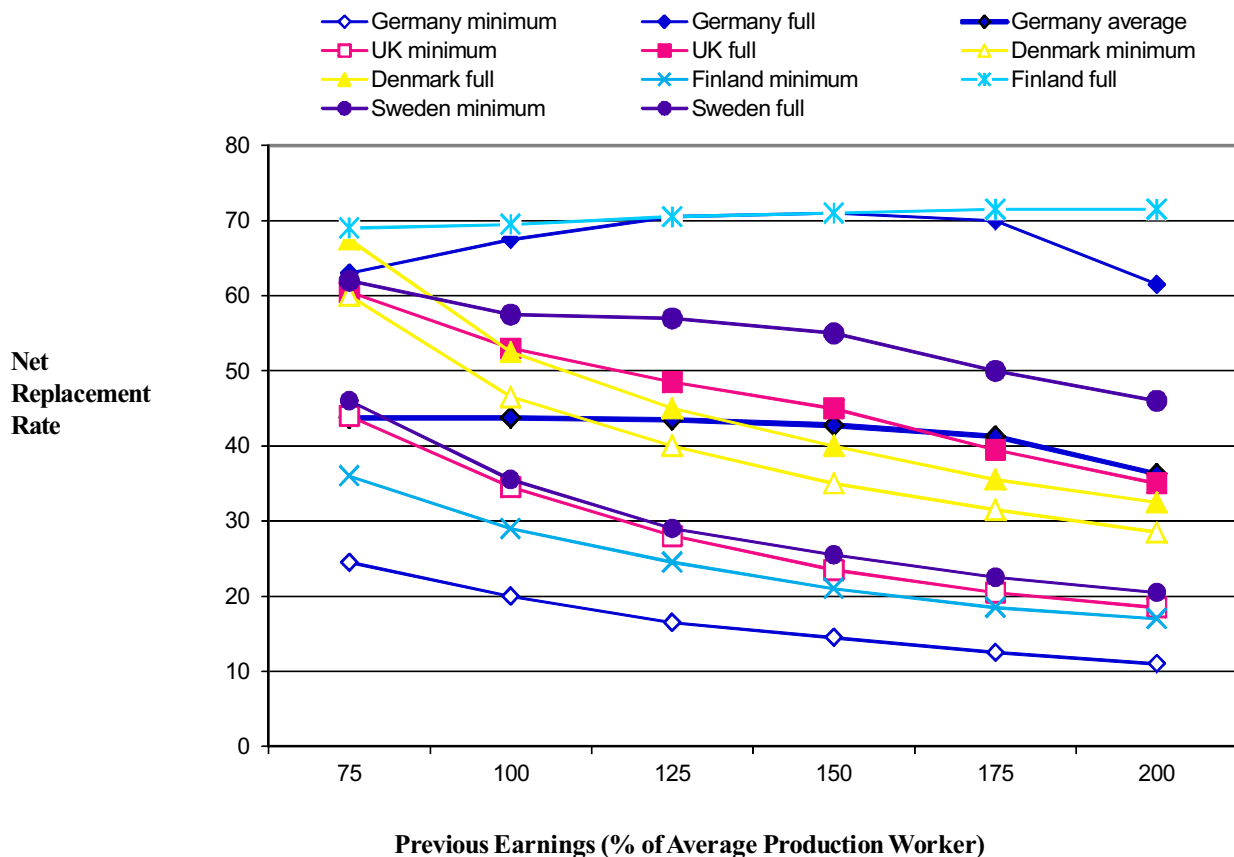
Compared with single-case indicators, composite indicators are more robust to differences across typical cases. An example is the OECD summary measure of benefit entitlement for unemployment benefits that is made up by the average of the gross unemployment benefit replacement rates for two earnings levels, three family situations, and three durations of unemployment (OECD 2007a). Distinct, but perhaps not representative, replacement rates are made less distinct in this average of 18 typical cases.

The risk, however, is that the composite indicators provide a mirror image of a fictive world, an image that is misleading for theoretical advances and policy advocacy. Figure 4 shows the situation of labor market insiders with full pensions and outsiders with minimum pensions.

A good example is Germany at the earnings level of the average production worker. The German minimum pension is the smallest for all countries, whereas the size of the German full pension is only surpassed by a hair by the Finnish full pension. However, the average of the German minimum and standard pension at this earnings level is situated in the middle, i.e., above the level of the minimum pensions in all countries, (except Denmark) and below the full pensions. The measure of the average pension in Germany thus says little about the dual nature of the social rights and the pension system in Germany, and it misinforms us about the cross-national differences.

Indeed, comparing minimum and full pensions demonstrates important points about social rights and the different nature of public old-age pension systems in the five countries. The gap between minimum and full pensions is largest in Finland and Germany and smallest in Denmark. This result indicates how the public pension systems in Finland and Germany follow the quid pro quo strategy, providing the most generous pensions of all countries for labor market insiders and the least generous for outsiders. In contrast, the Danish public pension system is characterized by giving everybody nearly the same pension, almost independent of labor market or contribution record. Social rights in Finland and Germany are largely based on work history, in Denmark on residence. Perhaps surprisingly, the gaps between Swedish minimum and full pensions are not negligible, especially for middle- and high- income groups, and larger than the gaps between similar British pensions.

Figure 4: Full and Minimum Old-Age Pensions in Five European Countries as a Percentage of Previous Earnings, Single Person, 2006



Concluding Remarks

I have shown that the practice of taking aggregate large-N measures into medium N-studies has its problems. Aggregate measures like social expenditures have problems in establishing equivalence of meaning across cases when they do not include the tax system and private social benefits. Also, social expenditure measures fail to give clear distinctions of welfare state types and wrongly assume that countries with similar expenditure levels also have similar welfare state types. Single-case indicators risk flaws in making inferences from the specific case to other cases. Group averages risk masking intra-group differences. Composite indexes mix apples and oranges when they encompass information on benefit coverage and generosity and tell us little about either coverage or generosity.

To better capture the complex, configurational nature of welfare state types, scholars may benefit from rethinking their use of large-N measures in medium-N studies. Given the need for multiple measures, the practice of making aggregate measures should be replaced by a broad range of measures. Such a set of multiple measures should capture what is of central importance. For example, as family types are becoming more diverse and the male breadwinner model is less dominant, we need to investigate the situation of not only the dominant family type—increasingly dual-earner—but also the other fam-

ily types. Similarly, when private social benefits are becoming of more importance, scholars may be ill-advised to stick with the study of only public social benefits.

References

- Adcock, Robert and David Collier. 2001. "Measurement Validity: A Shared Standard for Qualitative and Quantitative Research." *American Political Science Review* 95:3, 529–546.
- Adema, Willem and Maxime Ladaïque. 2009. "How Expensive is the Welfare State? Gross and Net Indicators in the OECD Social Expenditure Database (SOCX)." *OECD Social, Employment and Migration Working Papers*, No. 92, OECD Publishing.
- Esping-Andersen, Gøsta. 1990. *The Three Worlds of Welfare Capitalism*. Cambridge: Polity Press.
- Esping-Andersen, Gøsta. 2009. *The Incomplete Revolution. Adapting Welfare States to Women's New Roles*. Princeton: Princeton University Press.
- Korpi, Walter and Joakim Palme. 2007. *The Social Citizenship Indicator Program (SCIP)*. Swedish Institute for Social Research, Stockholm University.
- OECD. 2011. Social Expenditure Database via <http://www.oecd.org/els/social/expenditure>, retrieved 6 October 2011.
- Ragin, Charles. 2008. *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. Chicago: University of Chicago Press.
- Scruggs, Lyle. 2004. *Welfare State Entitlements Project Codebook*, version 1.1.
- Scruggs, Lyle. 2006. "The Generosity of Social Insurance, 1971–2002." *Oxford Review of Economic Policy* 22:3, 349–364.

Step Away from the Pool

Lane Kenworthy

University of Arizona

lane.kenworthy@gmail.com

Quantitative macrocomparative analysis of the rich long-standing-democratic nations—“medium-N analysis”—is dominated by pooled time-series cross-section regression. I estimate that more than 90% of the medium-N papers I read in journals and journal submissions use pooled regression.¹

Quantitative data on many of the institutions, policies, and socioeconomic outcomes studied by comparativists first became available for more than a handful of countries in the 1970s. For a while quantitative macrocomparative research consisted mainly of cross-sectional analysis of single-point-in-time data or period averages. By the 1990s reasonably lengthy time series existed and analysts began examining over-time patterns. Pooling over time and across countries helped alleviate what had long been considered the achilles heel of cross-sectional comparative analysis: the small-N problem, which limits the number of control variables that can be included in a regression model. Pooled regression became the tool of choice. It has remained so for two decades.

In my view, that’s unfortunate.

Better Apart

Aside from maximizing the number of observations, in many instances there is little reason to prefer pooling. Patterns of association across nations frequently differ from those over time. So why combine them? The default should instead be to examine them separately. This point has been made before—indeed, for quite a while now—so I won’t dwell on it (Griffin et al. 1986; Kittel 1999, 2008; Kenworthy 2007, 2009; Shalev 2007).

But that leaves us with only 30 or so years, or only 20 or so countries. What about omitted-variable bias due to the small N? The degree of worry about this among comparativists is far out of proportion to its actual danger. Overspecified models are just as likely to mislead as under-specified ones (Liebersohn 1985; Achen 2005), and the ability to throw in lots of controls tends to reduce the thought we put into our choice of models (Achen 2002).

Focus on Cross-Country Variation in Long-Run Change

Our aim should be to analyze changes rather than levels. Examining change gives us a better chance at identifying true causation.

Suppose we have data for 20 nations with annual observations from 1979 to 2007 (both years are business-cycle peaks). And suppose our interest is in the impact of a particular hypothesized cause. What type of analysis should we do?

The ideal is a differences-in-differences design, in which we compare changes over time across the countries. This is the closest approximation to an experiment. In an experiment we have two (or more) groups. We observe them at time 1. We

administer the “treatment” to some observations. Then we compare the degree of change in both (or all, if there are more than two) groups. With observational data we don’t administer the treatment, but in other respects the process is similar. Especially useful is that this design takes constant country-specific differences (“country fixed effects”) out of play.

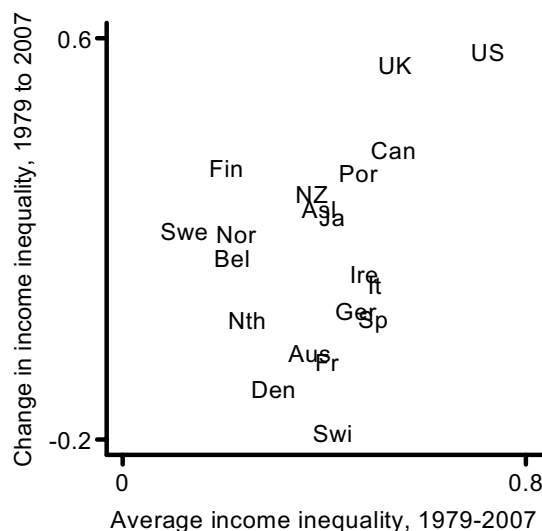
There remains plenty to worry about: selection into the treatment and control groups, reverse causality, country differences in other things that change, getting the lag right, variation in effects across subperiods. But this is a good design when appropriate.

When is it appropriate? When we have significant over-time change in the hypothesized cause, when there is variation across countries in that change, and when the change is mainly unidirectional rather than back-and-forth.

Doesn’t this throw out a lot of interesting year-to-year variation? It does throw out that variation, but in many instances that variation has little or no bearing on our research question.

Here’s an example. In recent years social scientists have grown increasingly interested in the effect of income inequality on socioeconomic outcomes such as health, crime, trust, and educational attainment (Burtless and Jencks 2003; Wilkinson and Pickett 2009). Inequality has risen significantly in some affluent nations. Figure 1 shows variation in income inequality change (vertical axis) and variation in income inequality levels (horizontal axis) for 20 countries. The range of values is the same on both axes.

Figure 1: Levels of Income Inequality and Changes in Income Inequality, 20 Countries, 1979–2007



Note: The income inequality measure combines the Gini coefficient for the bottom 99% of the income distribution and the top 1%’s share of income, with each rescaled to vary from 0 (smallest observation during the period) to 1 (largest). The same range of values, 0.8, is used on both axes. The change measure on the vertical axis is calculated as 2007 level minus 1979 level. Source: Jencks and Kenworthy, forthcoming.

Income inequality nicely fits the criteria for a differences-in-differences analysis. First, the amount of over-time change is substantial. Second, it varies a good bit across the countries. In fact, the cross-country variation in change between 1979 and 2007 is greater than in average levels over that period. Third, much of the change is unidirectional, as revealed by time plots for each of the countries (not shown here).

We could pool the data, using country-years or country-periods as observations, and do a pooled regression. But why? If a change in income inequality affects life expectancy, college completion, or trust, that causal process is likely to play out relatively slowly. Better, in my view, to use a simple differences-in-differences design, regressing change (first difference) in the outcome over the whole of the period on change in inequality.

This approach wasn't practical 20 or in some instances even ten years ago. The time series data didn't cover a long enough span of time. Now, for at least some interesting research questions, they do.

But Not Always

Next, consider an equally interesting and important question: What is the impact of tax levels on economic growth? Researchers have understandably—given the considerations noted earlier—tried to exploit the now fairly lengthy time series on tax revenues in OECD nations to get a handle on this question. Some analyses focus on the over-time variation within countries, some pool the over-time with the across, and some examine cross-country differences in over-time patterns.

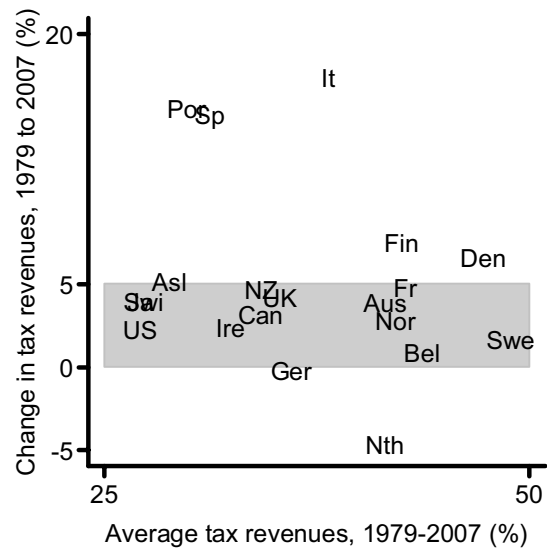
A common tendency here, as in much macrocomparative research, has been to use yearly data. That makes little sense. The standard measure of taxation level is tax revenues as a share of GDP. This tends to move in a predictable and consistent way within business cycles, rising during growth phases and declining during recessions. This movement is of no relevance to the question of taxation's effect on economic growth. It's noise rather than signal. Our focus should be on the long run, not on fluctuations within business cycles.

How much variation is there in long-run change? Figure 2 does the same thing with tax revenues as a share of GDP that Figure 1 did with income inequality: On the horizontal axis is each country's average level during the period 1979–2007 and on the vertical axis is the amount of change between 1979 and 2007. The range of values is the same on both axes.

Here we see limited cross-country variation in change. More than two-thirds of the countries—those highlighted by the shaded horizontal band—experienced changes (increases) in tax revenues of less than five percentage points over these three decades. There was significant change in only four countries: Portugal, Spain, Italy, and the Netherlands. In a differences-in-differences analysis, this small set of countries is likely to drive the findings.

There is much more variation in average levels of taxation, shown on the horizontal axis. This looks to be a more promising source of information about the causal effect of tax levels on economic growth. It calls for a straightforward cross-sectional analysis of levels. We can regress average economic

Figure 2: Levels of Taxation and Changes in Taxation, 20 Countries, 1979–2007



Note: Tax revenues as a share of GDP. The same range of values, 25 percentage points, is used on both axes. The change measure on the vertical axis is calculated as 2007 level minus 1979 level. Data source: OECD.

growth over 1979–2007 on average tax revenues over that period.

But doesn't that put us back in the small-N situation lamented by affluent-nation comparativists a generation ago? Yes, but the proper reason for lament was not the small N per se. It was the lack of sufficient data to know whether the analysis should instead focus on changes.

So what should we do? Everything we normally would do. Try some controls. Check for reverse causality. Consider influential cases. Look for interactions. Use case knowledge to ponder country fixed effects. Supplement the aggregate analysis with in-depth examination of a few countries.

Taxation isn't exceptional. There are lots of hypothesized causes in comparative politics and comparative political economy that have changed little over a fairly lengthy period of time: corporatism, cumulative left government, social policy generosity, welfare-state regime type, and health-care system, among others.

Summary

For the past two decades, a large amount of quantitative macrocomparative research on the world's rich nations has used pooled time-series cross-section regression with annual observations. I suspect this owes mainly to first-mover advantages and path dependence. In some instances, perhaps many, it's counterproductive. We almost always should separate the variation over time within countries from the variation in levels across countries.

Often we should focus on cross-country variation in long-run change. A differences-in-differences analysis is a good way to do this.

Sometimes there is little cross-country variation in

changes, but quite a bit in levels. In such instances a cross-sectional analysis of levels is appropriate. It's rare to see this in published research these days. It probably strikes many scholars as primitive and archaic—a relic of a bygone era. That's a mistaken and unhelpful view.

Note

¹ I'm here referring only to papers that use solely country-level data. A growing number combine country and individual data and thus use a multi-level technique.

References

- Achen, Christopher H. 2002. "Toward a New Political Methodology: Microfoundations and ART." *Annual Review of Political Science* 5, 423–450.
- Achen, Christopher H. 2005. "Let's Put Garbage-Can Regressions and Garbage-Can Probits Where They Belong." *Conflict Management and Peace Science* 22, 327–339.
- Burtless, Gary and Christopher Jencks. 2003. "American Inequality and Its Consequences." In *Agenda for the Nation*. Henry Aaron, Pietro S. Nivola, and James M. Lindsay, eds. (Washington, DC: Brookings Institution), 61–108.
- Griffin, Larry J., Pamela Barnhouse Walters, Phillip O'Connell, and Edward Moor. 1986. "Methodological Innovations in the Analysis of Welfare-State Development: Pooling Cross Sections and Time Series." In *Futures for the Welfare State*. Norman Furniss, ed. (Bloomington: Indiana University Press), 101–138.
- Jencks, Christopher and Lane Kenworthy. Forthcoming. *Should We Worry About Inequality?* New Haven and New York: Yale University Press and Russell Sage Foundation.
- Kenworthy, Lane. 2007. "Toward Improved Use of Regression in Macrocomparative Analysis." *Comparative Social Research* 24, 343–350.
- Kenworthy, Lane. 2009. "The Effect of Public Opinion on Social Policy Generosity." *Socio-Economic Review* 7, 727–740.
- Kittel, Bernhard. 1999. "Sense and Sensitivity in Pooled Analysis of Political Data." *European Journal of Political Research* 35, 225–253.
- Kittel, Bernhard. 2008. "Statistical Narratives and the Properties of Macro-Level Variables: Labor Market Institutions and Employment Performance in Macrocomparative Research." In *Method and Substance in Macrocomparative Research*. Lane Kenworthy and Alexander Hicks, eds. (Basingstoke, UK: Palgrave Macmillan), 29–66.
- Lieberson, Stanley. 1985. *Making It Count*. Berkeley: University of California Press.
- Shalev, Michael. 2007. "Limits and Alternatives to Multiple Regression in Comparative Research." *Comparative Social Research* 24, 261–308.
- Wilkinson, Richard and Kate Pickett. 2009. *The Spirit Level: Why Greater Equality Makes Societies Stronger*. New York: Bloomsbury Press.

Fast-Moving and Slowing-Moving Factors: Taking Apart Time-Series Cross-Sectional Analysis

Gary Goertz

University of Arizona
ggoertz@u.arizona.edu

In this essay I offer reflections on some basic methodological issues that all those in the field of advanced industrial-democratic societies face.¹ The scope of this research is typically limited to advanced industrial-democratic countries which number about 20–35 depending on the period. Time-series data is often available, depending on the country, for about 30–40 years (i.e., 1970s–2000s). This time-series cross-sectional situation forms one of two or three standard types, which also include many cross-sections and few time periods (e.g., panel survey data) or many cross-sections and many years (e.g., international conflict data which has many dyads, i.e., 20,000+, and many years, often 1816–2000). I also draw on the international conflict literature which faces many of the same problems and has discussed the same basic problem.

The time-series cross-sectional data of comparative political economy faces many serious problems. There are various common solutions, including fixed effects for cross-sectional issues and lagged dependent variables and panel-corrected standard errors for temporal ones (see Beck 2001; Plümper, Troeger, and Manow 2005; and Kittel and Winner 2005 for good discussions in a comparative political economy context).

Inspired by Pierson (2004, chapter 3), we can think of the methodological issues that this field faces in terms of slow-moving (including stationary) versus fast-moving variables, as described in Table 1. In Table 1 I give some examples where the first variable is the independent variable and the second is the dependent.

Another way to think about fast versus slow moving is in terms of the location of the variance in the data. Fast-moving means more variation, while slow-moving means less. The extreme of slow-moving is constant, which is of course no variance at all. The extreme of fast-moving is that the data vary from one extreme to another every year. This provides a specific means for placing a given X – Y pair in one of the cells of Table 1.

My general proposition is that each of these cells has significant, but often different, methodological challenges. I focus mostly on the situation where one of the variables is slow moving, in part because I think the methodological challenges are more severe. As I discuss in more detail below, the standard methodology works best if the data (i.e., world) fits into the fast-fast cell. As such the most problematic cell for this methodology is the slow-slow one. The methodology for time-series cross-sectional is a "one size fits all." I think by default most scholars will analyze the data falling into these four cells in basically the same way. This is not to say that there is a

Table 1: Slow-Moving Versus Fast-Moving Variables

Inspired by Pierson (2004, Table 3.1)

		Independent Variable	
		Slow	Fast
Dependent Variable	Fast	Inequality–Left government	Democracy–civil war
	Slow (Rare Events) (Punctuated Equilibrium)	Inequality–Education Inequality–Civil War GDP/capita–Democracy	Left government– Social policy generosity

consensus on how to analyze time-series cross-sectional data; there is not, but these differences are not particularly tied to the distinctions made in Table 1.

My next move is to confront Table 1 with the typical time-series cross-sectional data. The basic statistical fact of social life is the following:

As the variables become slower moving more of the variation of the data is cross-sectional.

One can see this in the extreme case where both *X* and *Y* are very slow moving. The relatively large differences between countries will then account for much more of the total variance.

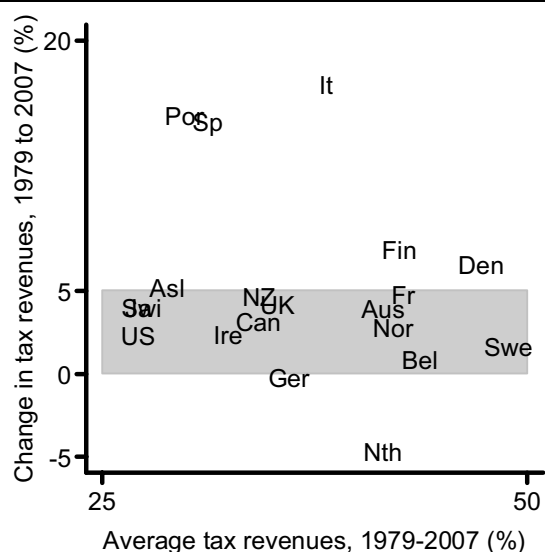
Many of the variables of most interest to comparative scholars fall into the category of slow-moving. Many of the basic, structural features of societies, economies, and government institutions change slowly almost all the time. There may be revolutionary or crisis moments, but the vast majority of the years in a time-series consist of small changes or no change to the status quo.

In Table 1 for the slow-moving dependent variable I have indicated a couple of hybrid categories. Punctuated equilibrium and rare events are general situations where the dependent variable can be considered to be fast moving but where most of the variation is cross-sectional. Hence they can pose the same methodological problems that a slow-moving dependent variable does, the predominance of cross-sectional variation over time-series. In fact, because the dependent variable in these types is often dichotomous, the methodological problems are more starkly posed than with slow-moving continuous variables.

Figure 1 illustrates the basic problem of slow-moving variables with a common comparative political economy variable, tax revenues as a share of GDP. Over a 30-year period there is little change in the tax revenues as a percentage of GDP. This figure just gives the difference between 1979 and 2007 so it masks oscillations in the data. But most of this oscillation is due to business cycles and as such is irrelevant for most research on tax revenues. The figure also illustrates that most of the variance is cross-sectional. There is quite a range of levels of taxation *between* countries, ranging from about 25 percent to over 50 percent.

**Figure 1: Variation Over Time:
Little Variation for Most Countries**

(Source: Lane Kenworthy)



Differences-in-Differences Methods

Differences-in-differences methods constitute a current standard for dealing with time-series cross-sectional data. It addresses exactly the problems of time-series cross-sectional data (e.g., Angrist and Pischke 2009, chapter 5). While the term is more common among economists, the issues and possible solutions are those addressed in the political science literature as well (e.g., Beck 2001; Plümper, Troeger, and Manow 2005; Kittel and Winner 2005). This is not one method but a family of methods and there are many hard choices to be made within this family and plenty of family squabbles. While it exceeds the scope of this essay, I think that the differences-in-differences makes concrete a long-standing position of many methodologists on the relative value of cross-sectional versus time-series data. This approach is useful in the context of this essay because it lays out in stark terms the issue that interests me.

Differences-in-differences methods typically are founded on the Neyman-Rubin causal inference framework. The basic starting point is a counterfactual contrasting the application

of the treatment (i.e., $X = 1$) versus the counterfactual of control ($X = 0$) for the individual subject i . In our case, the subject is the country. While one cannot perform experiments on countries, the idea is that causal contrasts should be *within subjects*. This leads to a central principle regarding time-series cross-sectional research:

Principle: Cross-sectional variation is a nuisance.

Ideally, all the causal contrasts should be based purely on within-country variation.

While I have stated the Principle in extreme form, there is a long history of methodologists expressing a clear preference for time-series variation over cross-sectional (see Ray 2003, 20–27 for a nice discussion). The difference-in-difference methodology just makes more explicit a long-standing preference of many methodologists.

If we consider the slow-slow cell of Table 1, we run up against a basic problem: The ideal is to use only variation within countries, but either (1) almost all the variation is cross-sectional or (2) the variables have a slow-moving consistent trend.

Fixed Effects

The clearest practical application of this Principle is the use of fixed effects for country. These are all nuisance variables, they are almost never reported in research articles, and I suspect that most scholars do not even look at the estimated fixed-effect parameters in any real sense.

The issue of slow-moving dependent variables appears in many ways in its starkest form with rare event–dichotomous dependent variables. With dichotomous dependent variables you have a clear contrast between change and no change, so the lack of time-series variation can appear more clearly. If the dependent variable is a rare event, that potentially means that for many cross-sections it does not occur at all; hence no variation on Y over time for those cross-sections. Many cross-sections may have only one event, a little variation but still not much at all.

Green, Kim, and Yoon (2001) proposed using fixed effects for time-series cross-sectional international conflict data. The symposium on that article illustrates the problems I think occur very typically with slow-moving continuous variables. The typical international conflict dataset is a rare event (militarized conflict is rare among dyads). Because the cross-section is the dyad, we have a large N of cross-sections. Given that there are less than 2,000 conflicts total, this means that many cross-sections have no variation on Y (i.e., no conflicts): “It turns out that only 198 of the dyads have any longitudinal variation in the dependent variable. The remaining 2,877 dyads are all zeros and have no effect on the regression results” (Green, Kim, and Yoon 2001: 455).

Almost all of the symposium regarding this article (with contributions by Oneal and Russett, Beck and Katz, and King) deal with the fact that a tremendous amount of information about conflict is being discarded by this approach. Virtually by definition this information is cross-sectional in nature. As Beck and Katz nicely put it in their contribution: Is one throw-

ing the baby out with the bath water? Their recommendation is clear: “We show that the use of fixed effects is clearly a bad idea for the binary dependent variable case” (Beck and Katz 2001: 488). They note exactly the problem that motivates this essay: “Relatively few interesting independent variables are temporally constant, although many are almost constant. These variables, like democracy, that vary little from year to year, are highly co-linear with the fixed effects. It is quite likely, then, that the use of fixed effects will yield odd estimates of coefficients for variables like democracy, since the effect of democracy is then ‘controlled’ for the fixed effects” (Beck and Katz 2001: 492–493).

I suggest that all research that fits into the the slow-slow cell of Table 1 has the same problems with fixed effects that are discussed by contributors to the Green et al. symposium. The dichotomous variable–rare events dependent variable just is a very clear example of the issues at stake for all studies in that cell.

A First Step Forward: Know Your Cross-Sectional Versus Time-Series Variation

One of the first obvious suggestions based on Table 1 is to actually investigate the extent to which key variables are slow or fast moving. It is quite possible, and not hard, to calculate the within-country variation versus the between-country variation on all key variables.

I think it should become a standard part of any methodological discussion that this basic information be given. In many ways this information is significantly more critical than the standard descriptive statistics one sees in many journal articles. This information then helps identify the specific kinds of theoretical and methodological problems the research faces.

Table 2 illustrates what this looks like for some standard comparative political economy variables, for some 20 usual OECD countries for a standard time-series period. The table is ordered by decreasing ratio of between-to-within variance. As a rule for this representative dataset, there is two to three times more variation between countries than over time.

If we look at the those variables that have relatively more within variation, we see that most of the them are slow moving, but with a consistent over-time trend. For example, most of these industrialized democracies have pretty consistent GDP/capita growth records over this 30-year period. Clearly percentage elderly is slow moving as these societies are gradually aging. The same is true for infant mortality, which has generally been going down over this period. Analyzing these variables with differences-in-differences is going to be problematic because they change so little from year to year and one has a clear trend line for most countries.

The only fast-moving variable in Table 2 is Left government. Here is where we see that within variation really dominates and we have major shifts in a short periods of time.

In short, most of the information we have for studying the causes and effects of key variables is cross-sectional.

**Table 2: Cross-Sectional Versus Time-Series Variation:
Common Political Economy Variables: 1979–2007, 20 OECD Countries**
(Source: Lane Kenworthy)

Variable	Between	Within	Between/Within
Decommodification	5.6	1.6	3.5
Poverty	4.1	1.4	2.9
Unionization	20	7	2.9
Trade	32	11	2.9
Employment	7.3	3.6	2.0
Gov. Social Expenditures	4.6	2.4	1.9
Education	1.7	1.0	1.7
Gov. Expenditures	7.3	4.2	1.7
Elderly Population	1.9	1.5	1.3
GDP/capita	39	43	0.9
Infant Mortality	2.5	2.7	0.9
Left Government	20	33	0.6

Causal Inference: Cross-Sectional Versus Temporal Variation

The differences-in-differences Principle is an ideal. In practice the parameter estimate for a given independent variable uses both time-series and cross-sectional information. An implicit hypothesis underlies most of this research:

Causal inference assumption: Time-series and cross-sectional information lead to the same inference and approximately the same parameter estimates.

There is no particular reason why this is likely to be true, as was emphasized by Lieberman and Hansen over 30 years ago (1974).

The standard statistical analysis of time-series cross-sectional data uses a complex mixture of cross-sectional and time-series variation. Rarely is the researcher really even aware of the character of this mixture and almost never is the mixture a topic of analysis. The implicit assumption is that cross-sectional evidence agrees with the time-series evidence. This assumption is a very important one when the variables are fast moving, and where there is typically both significant cross-sectional and significant time-series variation.

Ray (2003) illustrates this with international conflict data, using basically the same data as in the Green et al. symposium. He contrasts various standard independent variables, some of which are fast moving and some are slow moving. He breaks a given variable into its two variance components over time and over space (using deviations from the mean, basically a fixed-effects approach).

Table 3 reproduces his results for one of the most slowest-moving independent variables, geographical contiguity (which is almost constant within dyads), and dyadic trade, which is one of the fastest-moving independent variables he analyzes. As one can see, the conclusions about the relationship between trade and conflict (hypothesized to be negative by the liberal peace literature) changes completely depending on whether one relies on the time-series information or the cross-sectional variation.

The contiguity variable does not have this problem because virtually all the variation is cross-sectional. There is no potential conflict because the variable is extremely slow moving, and virtually constant for almost all dyads.

Also problematic is how scholars implicitly convert inferences based largely on cross-sectional data to cross-time inferences. For example, the classic literature on the relationship between GDP/capita and democracy is based on mostly cross-sectional evidence (GDP/capita is slow moving and democracy fits the punctuated equilibrium pattern). This cross-sectional evidence is then naturally converted into hypotheses about changes over time: “If other countries become as rich as the economically advanced nations, it is highly probable that they will become political democracies” (O’Donnell 1973: 3). I suspect that if one looks at the interpretation of parameters (e.g., logit as well as OLS) they imply over-time changes that are implausible: They are turning cross-sectional variation into temporal.

On the other hand, there must be some relationship between the over-time and over-space parameter estimates. I think that figuring this out should be a large part of the research agenda for time-series cross-sectional analyses.

Qualitative Methodology

If the differences-in-differences philosophy rejects cross-sectional analyses, qualitative methodology has historically embraced it using the comparisons between countries as central to understanding social and political phenomena. A very large percentage of the classics of comparative politics involve intensive comparisons of a small number of countries.

However, the same issues I discussed in the previous section arise for qualitative comparative work as well. Given the temporal nature of case studies, these issues will often be raised, but an explicit asking of the question about cross-sectional versus time-series conclusions in individual cases would certainly be useful.

Looking at overall statistics on variance often masks important features of individual cases. When I say a variable is

Table 3: Issues with Slow-Moving Variables in Time-Series Cross-Sectional Data
(Source: Ray 2003, table 1; see also Zorn 2001)

Variable	Parameter Estimate	Z-score	Significance
Contiguity (slow moving)	3.5	47.7	.000
Trade (cross-sectional)	-34.6	-4.2	.000
Trade (time-series)	48.9	4.9	.000

Note: Dependent Variable is "Militarized Conflict"

slow-moving, that typically means it is slow-moving for the vast majority of cases. However, it is quite possible that there are a few fast-moving countries in the data, as illustrated in Figure 1. The mass of slow-moving countries can hide these cases.

Social science might follow medicine, which has a long tradition of publishing case studies of "interesting" patients. For example, Oliver Sacks has made an extremely successful career of using very special cases to try to understand how the brain works. One can potentially gain a lot of insight by focusing on the few cases that seem to be clear outliers. These kinds of case studies will almost certainly involve an implicit, but ideally explicit, contrast with other cases which seem to exemplify the normal kinds of variation.

Conclusion

Ray in his discussion of these issues in international relations (and other subfields as well) notes that:

it so happens that some of the most robust, and theoretically important relationships within international politics and within political science, generally speaking, happen to be cross-sectional rather than longitudinal or over time relationships. This is not the place to speculate in detail about why this might be the case. But one obvious reason is that many important variables are time invariant. (Ray 2003: 22)

A main point in this essay is that with slow-moving variables one needs to embrace cross-sectional variation. Often that is the only information we really have. So we should reject philosophies that deny the usefulness of cross-sectional data analysis for social science.

At the same time, one cannot just assume that the results of a cross-sectional analysis apply to causal inference within individual countries. The Ray example above as well as the classic Lieberman and Hansen (1974) article show it is quite possible for the results of the two to differ. The statistical literature clearly separates the cross-sectional from temporal problems and I think that qualitative methods could well do the same in their own context. Process tracing is the qualitative methodology for within-case over-time causal analysis. One should explicitly ask about the extent to which process tracing conclusions fit with cross-national comparisons. The goal is to give an account which makes sense of both cross-sectional and time-series results.

Note

¹ I would like to thank Jonus Pontusson for several brainstorming sessions; Lane Kenworthy for many useful comments, for much of the data, and for the permission to use Figure 1; and Patrick Emmenegger for comments on an earlier draft.

References

- Angrist, Joshua D. and Jörn Steffen Pischke. 2009. *Mostly Harmless Econometrics*. Princeton: Princeton University Press.
- Beck, Nathaniel. 2001. "Time-Series-Cross-Section Data: What have We Learned in the Past Few Years?" *Annual Review of Political Science* 4, 271–293.
- Beck, Nathaniel and Jonathan Katz. 2001. "Throwing Out the Baby with the Bath Water: A Comment on Green, Kim, and Yoon." *International Organization* 55:2, 487–495.
- Green, Donald P., Soo Yeon Kim, and David H. Yoon. 2001. "Dirty Pool." *International Organization* 55:2, 441–468.
- Kittel, Bernhard and Hannes Winner. 2005. "How Reliable is Pooled Analysis in Political Economy? The Globalization-Welfare State Nexus Revisited." *European Journal of Political Research* 44, 269–293.
- Lieberman, Stanley and Lynn K. Hansen. 1974. "National Development, Mother Tongue Diversity, and Comparative Study of Nations." *American Sociological Review* 39, 523–541.
- O'Donnell, Guillermo A. 1973. *Modernization and Bureaucratic Authoritarianism: Studies in South American Politics*. Berkeley: Institute of International Studies.
- Pierson, Paul. 2004. *Politics in Time: History, Institutions, and Social Analysis*. Princeton: Princeton University Press.
- Plümper, Thomas, Vera E. Troeger, and Philip Manow. 2005. "Panel Data Analysis in Comparative Politics: Linking Method to Theory." *European Journal of Political Research* 44, 327–354.
- Ray, James Lee. 2003. "Explaining Interstate Conflict and War: What Should be Controlled For?" *Conflict Management and Peace Science* 20, 1–30.
- Zorn, Christopher. 2001. "Estimating Between- and Within-Cluster Covariate Effects, with an Application to Models of International Disputes." *International Interactions* 27, 433–445.

*Political Science Seen Through
Historians' Looking Glass:
The Case of Electoral System Choice*

Patrick Emmenegger

University of Southern Denmark
emm@sam.sdu.dk

Klaus Petersen

University of Southern Denmark
klaus.petersen@hist.sdu.dk

The methodological challenges of medium-N research have triggered a vibrant debate in the political science community. The contributions to this symposium lay testimony to this. In this short article, we want to add a new, previously neglected dimension to this methodological debate. We explore how area experts, in this case historians specialized in the areas of political, economic, and labor history in the period 1890 to 1939, perceive state-of-the-art comparative political economy work. We use the debate on electoral system choice in the decades before World War II as our working example (Alesina and Glaeser 2004; Blais et al. 2005; Boix 1999, 2010; Cusack et al. 2007, 2010; Kreuzer 2010).¹ There are a number of reasons for this choice: First, electoral system choice is one of the most important topics in contemporary political science research. Sartori (1968: 273, quoted in Lijphart 1994: 139) calls electoral systems “the most specific manipulative instrument in politics.” In a similar vein, electoral system choice is a highly relevant topic in historical research. Thus, both historians and political scientists are working on electoral system choice in the decades before World War II.

Second, the authors involved in this debate are among the leading researchers in comparative political economy. In addition, the methodological approaches chosen by the participants in this debate are quite typical of macrocomparative political economy research.

Finally and most importantly, the debate is characterized by widespread disagreements and repeated rejections of evidence. The four arguments put forward make theoretically contradictory predictions. For instance, while Blais et al. (2005) and Cusack et al. (2007, 2010) emphasize societal consensus on the issue of electoral system choice, Alesina and Glaeser (2004) and Boix (1999, 2010) stress political conflict. In a similar vein, while Alesina and Glaeser (2004) and Blais et al. (2005) identify forces of democratization (the labor movement and the spread of democratic ideas respectively) as the drivers of electoral reform, Boix (1999, 2010) and Cusack et al. (2007, 2010) highlight the role of old political elites that attempt to secure their positions of power in the political system through electoral reform. Methodologically, the contributors to this debate cannot agree on the positions of political parties on electoral system choice (Cusack et al. 2007: 379–382; Kreuzer 2010: 376; Boix 2010: 407), the extent of conflict in these countries during the (non-)adoption of proportional representation (Cusack et

al. 2010: 396; Kreuzer 2010: 381; Boix 2010: 407), or the extent to which countries were characterized by large skill-based export sectors (Cusack et al. 2007: 385; Kreuzer 2010: 372). As a result, Cusack et al. (2007: 386) find Boix's (1999) socialist threat variable to be a bad predictor of the effective electoral threshold, while Kreuzer (2010: 379) finds it very robust. In turn, Boix (2010: 410) finds Cusack et al.'s coordination index (2007) to be an insignificant predictor of party support for proportional representation, while in the account of Cusack et al. (2007: 386) the coordination index can explain 83 percent of the variation of the dependent variable (the effective electoral threshold).

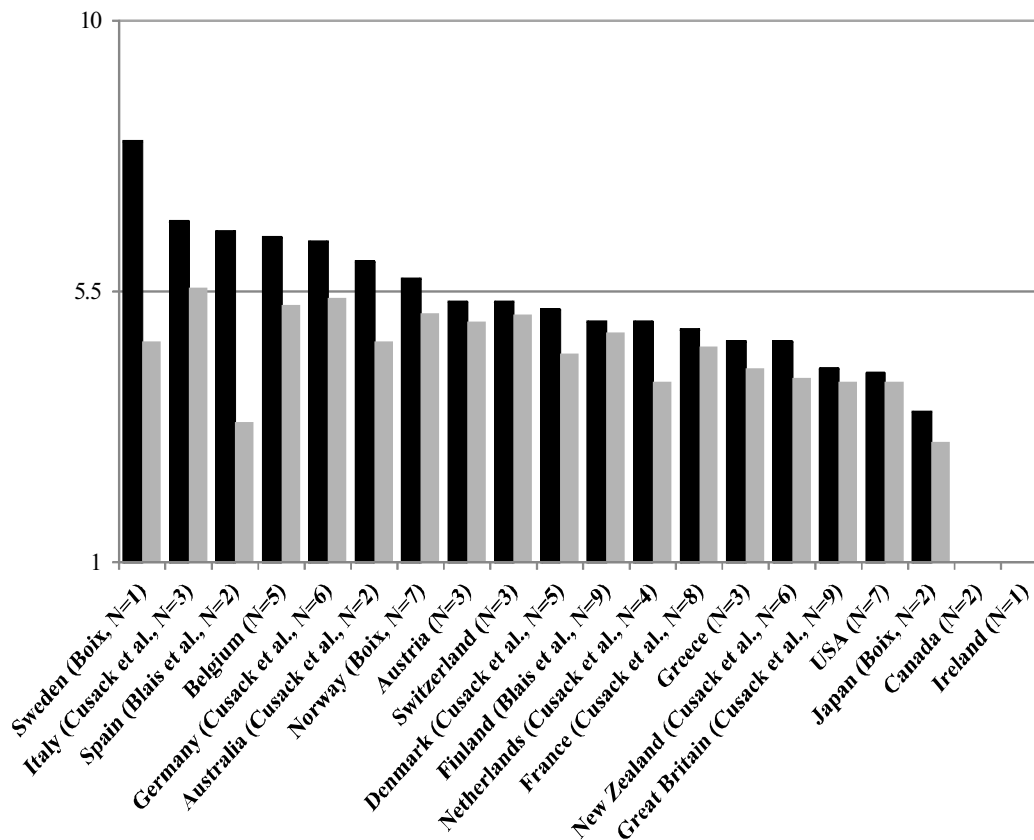
In a discussion of Boix (1999) and Cusack et al. (2007), Kreuzer (2010) believes he has found a reason for this situation. He argues, “political scientists commonly draw on history but often do not read actual historians carefully” (Kreuzer 2010: 369). Moreover, he argues, “it would be beneficial to first do the more nuts-and-bolts work of using historical knowledge to improve the quantitative study of institutions origins” (Kreuzer 2010: 385). He concludes with the call to “take not only history, but also historians seriously” (Kreuzer 2010: 383).

Taking Historians Seriously: An Expert Survey

In this short article, we want to answer Kreuzer's call. For this purpose, we have conducted an expert survey among historical experts specialized in political, economic, and labor history in the period 1890 to 1939 in 20 countries. We compiled the list of 300 experts (historians and historically oriented social scientists) by contacting leading national and international history journals and asking them to nominate experts to participate in the survey. Subsequently, we contacted these historical experts and asked them to participate in an online survey. The survey consisted of four parts. In the first part, we asked for the participants' academic background, countries and periods of expertise, and the extent of collaboration with social scientists (defined as political scientists and sociologists). Subsequently, we asked them to assess on a very general level the usage of history by social scientists (the statements were inspired by the discussions in Elman and Elman [2001] and Thies [2002]). The third part of the survey was introduced by a short description of the debate on electoral system change and the summaries (max. 400 words) of the arguments by Blais et al. (2005), Boix (1999, 2010), and Cusack et al. (2007, 2010).² In the case of the arguments by Boix (1999, 2010) and Cusack et al. (2007, 2010) we incorporated their recent revisions. The participants in the survey were then asked to rate the three arguments from 1 (does not at all explain the presence or absence of electoral system change) to 10 (fully explains the presence or absence of electoral system change). In the last part of the survey, the participants were asked a second battery of general questions, centering on issues such as the need for generalizations and the usage of quantitative approaches. Throughout the survey, participants had the opportunity to complement their answers with written comments. (A detailed description of the survey and the short summaries of the arguments are available at: www.sdu.dk/welfare.)

Figure 1 displays the main results of the expert survey. The black bars in Figure 1 show for each country the average

Figure 1: Evaluation of the Three Arguments About Electoral System Choice



Note: The figure shows the average score of the best-performing explanation (black bars) and the average score across all three explanations (gray bars). The best performer, where there is a single-best performer, is listed in the X-axis labels following the country name together with the number of observations.

evaluation of the best-performing explanation, while the gray bars show the average evaluation across all three explanations. Three lessons can be drawn from this data. First, the evaluation is overall rather negative. In the case of the best-performing explanation for each country (black bars), only 7 out of 20 cases score above the point of maximum ambiguity (5.5). In two cases (Canada and Ireland), all explanations received an average score of 1 (“does not explain at all the presence or absence of electoral system change”). In the case of the average performance across all three explanations (gray bars), only 1 out of 20 cases scores above the point of maximum ambiguity (Italy at 5.56).

Second, the historical experts consider Cusack et al.’s argument (2007, 2010) most convincing. It received the highest average score in 8 out of 20 cases (Australia, Denmark, France, Germany, Great Britain, Italy, the Netherlands, and New Zealand). In contrast, historical experts consider in three cases (Japan, Norway, and Sweden) and two cases (Finland and Spain) respectively Boix (1999, 2010) and Blais et al. (2005) to be the most convincing explanation. In the remaining seven cases, at least two explanations scored identically.

Finally and somewhat surprisingly, “pure historians” evaluated the three arguments more positively than historical experts who declared to have at least some social science back-

ground. The latter gave the three arguments an average score of 4.11; “pure historians” gave them 4.72. This difference is significantly different from zero at the 10 percent level.³ Thus, the negative evaluation is unlikely to be caused by a lack of familiarity with social science research.

How can we explain these results? At first sight, we might dismiss them by pointing to disciplinary differences, in particular the (often claimed) general reluctance on the part of historical experts to endorse quantitative approaches, generalization, and comparison (see Elman and Elman [2001] for a critical discussion of conventional conceptions). However, as Table 1 documents, these three factors are unlikely to explain the negative evaluation. After evaluating the three arguments about electoral system choice, the survey participants generally expressed high levels of support for generalizations and comparisons, and even some support for quantitative approaches. Thus, the negative assessment seems to be the result of other factors.

As Table 2 documents, two other factors seem to play a more important role. First, most historical experts are doubtful that we can always give clear and simple answers to questions about historical facts. This is important because especially in macro-quantitative political economy research we are generally dependent on being able to assign scores to every case

Table 1: Survey Results on Generalization, Comparison, and Quantitative Approaches (Percent)

Does it make sense to look for explanations that can be generalized to more than one country?		How much can we learn about our country of expertise by comparing it with other countries?		Can a quantitative approach contribute to the explanation of historical developments?	
Always	22.7	A lot	57.6	Always	3.0
Almost always	33.3	Quite a lot	33.3	Almost always	15.2
Sometimes	39.4	Occasionally	7.6	Sometimes	48.5
Almost never	3.0	Rarely	1.5	Almost never	25.8
Never	1.5	Never	0.0	Never	7.6
N	66	N	66	N	66

Note: The complete questions are: (1) “Does it make sense to look for explanations that can be generalized to more than one country (e.g., in the analysis of historical events or developments)?” (2) “How much can we learn about our country of expertise by comparing it with other countries and the development of our outcome of interest (e.g., electoral system change) in these countries?” (3) “Some social scientists advocate a quantitative approach to historical developments. These scholars argue that researchers should not focus on countries but rather on ‘variables.’ In particular, they argue that social scientists should develop indicators measuring the ‘spread of democracy’ in the neighboring countries (Blais et al.), the level of a ‘socialist threat’ to established parties (Boix), or the level of ‘coordination between different group interests’ (Cusack et al.) and test whether these variables are systematically associated with the outcome of interest (here electoral system choice). Can this quantitative approach contribute to the explanation of historical developments?”

Table 2: Survey Results on Historical Facts and Medium-N Approaches (Percent)

Are we able to give clear answers to questions about historical facts (such as consensus or conflict)?		Are (variable-oriented) medium-N approaches a reasonable way to study historical developments?	
Always	4.5	Always	6.1
Almost always	31.3	Almost always	12.1
Sometimes	58.2	Sometimes	12.1
Almost never	4.5	Almost never	47.0
Never	1.5	Never	22.7
N	66	N	66

Note: The complete questions are: (1) “Some of the social scientists (political scientists and sociologists) contributing to the debate on electoral system change in the period 1890–1939 disagree on historical facts. Some argue that there was conflict between parties on the issue of electoral system choice in a given country, some argue that there was none. Are we in general able to give clear answers to such questions for a given country?” (2) “The social scientists contributing to this debate work with approximately 20 countries in order to develop a general argument about electoral system choice. They argue that if their argument is correct for most of these 20 countries, then it should be considered a general explanation for all 20 countries. This means that even for countries for which the explanation is not correct, the argument should be considered an important part of the specific national explanation. Do you think that this is a reasonable way of studying historical developments?”

being analyzed. The level of consensus among political groups on the issue of electoral system choice is one of the key observable implications of the arguments put forward about electoral system choice. As mentioned above, Blais et al. (2005) and Cusack et al. (2007, 2010) emphasize societal consensus on this issue, while Alesina and Glaeser (2004) as well as Boix (1999, 2010) stress political conflict. Using this example, we asked the participants whether we are generally able to give clear answers to questions about issues such as the level of consensus among political parties at the climax of the debate about the adoption of proportional representation in the period 1890 to 1939. Somewhat surprisingly, 58 percent of the survey participants answered that it is only sometimes possible (see Table 2). Only about 36 percent answered that it is

always or almost always possible to give clear answers to such questions. This sheds a different light on the debate between Cusack et al. (2010, 2010b) and Kreuzer (2010, 2010b) on the coding of indicators. Some historical cases might simply defy a straightforward coding.

Second, historical experts are critical of what could be called variable-oriented medium-N approaches to studying historical developments (see Table 2). Variable-oriented medium-N approaches typically make generalizations across the whole sample (or the whole universe of cases) even though they cannot explain every case. This means that even for countries for which the explanation is not correct, the argument should be considered an important part of the specific national explanation. Almost 70 percent of the participants do not consider

**Table 3: Survey Results on General Perception of Social Scientists:
Ranked from 1 (Disagree) to 10 (Agree)**

	N	Mean	Std. Dev.
Social scientists' historical knowledge is generally very good	112	4.2	1.84
In general, social science research pays sufficient attention to the historical dimension of political and social developments	113	3.8	1.66
Social scientists are generally more interested in defending their theories than in analyzing actual historical developments	111	6.7	2.21
Social scientists often draw facts from historically grounded accounts and then use them for timeless propositions	105	6.7	1.96

this a reasonable way to study historical developments. Given the otherwise rather positive attitudes towards generalization, comparison, and quantitative approaches, it is quite likely that it is the focus on variables instead of cases that makes historical experts reject such approaches. In the comment sections, several experts expressed irritation about how an explanation could be considered a general explanation if it cannot explain *every case*.

This view is also confirmed by the survey results displayed in Table 3. These questions were asked before the empirical example was introduced. As is to be expected, historical experts are rather critical of social scientists' historical knowledge and think that social science research does not pay sufficient attention to the historical context of political and social developments. Potentially more damaging, they believe that social scientists are generally more interested in defending their theories than in analyzing actual historical developments and that social scientists often draw facts from historically grounded accounts and then use them for timeless propositions. For all four questions, the means are significantly different from the point of maximum ambiguity (5.5) at the 1-percent level. Interestingly, historical experts with some social science background are again more critical of social science research than are "pure historians."⁴

Thus, "taking historians seriously" (and asking them to evaluate our work) is a painful experience for political scientists. Overall, the evaluation is rather negative. It is particularly bewildering that the evaluation by historical experts with social science background is more negative than the evaluation by what we refer to as "pure historians." Therefore, the negative evaluation cannot be dismissed by pointing to historical experts' lack of familiarity with social science research. Nor can it be dismissed by pointing to disciplinary differences such as the (seemingly) general reluctance on the part of historical experts to endorse quantitative approaches, generalization, and comparison. The problem seems to lie elsewhere. Historical experts doubt that it is possible to provide clear answers to questions about historical facts (which we need to make indicators), they stress the context-dependence of historical events, and they reject variable-oriented approaches (that make generalizations across the whole sample even though they cannot explain all cases).

What can we learn from this?

Friendly competition between academic disciplines is normal and the historians' critical assessment has to be taken with a grain of salt. We might be equally critical of the historians' work. Nevertheless, if we take this assessment at face value, what can we learn from it? The two main messages of Figure 1 are certainly "equifinality" and "outliers." The historical experts have converged on different explanations for electoral system choice for the 20 countries covered by the survey. For instance, it seems as if the socialist threat was particularly important in Sweden, while Spain adopted proportional representation (PR) under the pressure to democratize and Italy's decision was influenced by economic considerations. Across the whole population of cases, the three arguments obtain similar average scores, but there is a lot of variation between the different countries. Overall, the expert survey lends support to claims that equifinality is in fact widespread. In addition, some countries seem to fit none of the accounts. As Figure 1 shows, none of the Canadian and Irish historical experts participating in the survey considered any of the three arguments helpful. One Irish survey participant wrote in the comment section that it is impossible to evaluate the three arguments given that the electoral system had been imposed by the United Kingdom. With regard to Canada, one participant wrote, "there are unique historical facts that explain the strong support for the first past the post system and the rejection of PR." Finland is another exceptional case. Becoming independent only in 1917, Finland immediately introduced universal suffrage and proportional representation because it had no room for disputes since "no party wanted to give the Russian government the chance of intervention."

It is certainly debatable to what extent standard methodological approaches, in this case correlation-based, are able to deal with equifinality and a considerable number of deviant cases. More damaging, given the importance of hypothesis testing in social science methodology, is the charge that social scientists are more interested in defending their theories rather than analyzing actual historical developments. This charge could also be formulated differently: To what extent are we choosing the toughest test possible for our hypotheses? All of the contributors to the debate on electoral system choice have used the electoral system in place as their dependent

References

variable (either as a dummy variable or in the form of the effective electoral threshold). Comparative political economy scholars, however, typically know the chosen electoral system in the countries under investigation. This choice of dependent variable implies that the new theory is not tested using new (previously unknown) data. Rather, the scores of the dependent variable are well known before the test. This is certainly not the toughest test possible.

Admittedly, our expert survey does not offer any methodological alternatives. The historical experts simply put their fingers on what they consider the weakest links in the social scientists' chain. Hence the first question for comparative social scientists is: Should we care about historians' opinions? There is certainly some truth in the stereotypical portraits of historians in love with particularities and political scientists looking for general laws of society (Elman and Elman 2001; Burke 2007). However, we believe that this is a matter of differences-in-degree rather than differences-in-kind. Without doubt, there are radical historicists rejecting any kind of generalization (a contacted historian described the very idea of an expert survey as "sociological humbug") and ultra-nomothetic political scientists determined to find general laws, but most scholars within both disciplines are pragmatic and open-minded. Certainly, the historians' negative assessment should not make us give up the ambition to create middle-range theories. After all, their focus on particularity and context comes at a price. In addition, there might be some things historians can learn from social scientists. However, we believe that we should take their criticism seriously.

Notes

¹ We focus on the period 1890 to 1939 because the first country to use a proportional representation system (PR) was Belgium in 1899 and because the onset of World War II with the accompanying collapse of several democracies using PR challenged the previous consensus on the democratic virtues of PR (Blais, Dobrzynska and Indridason 2005: 182, 186).

² We did not consider Alesina and Glaeser's argument (2004) in the expert survey because it is not as developed as the other arguments. In particular, it cannot explain why PR was sometimes introduced by the political right *against* the will of the political left and sometimes by the political left against the will of the political right. The argument is thus underspecified. However, we would like to emphasize that the political left is likely to have played a more active role than is acknowledged in the other three arguments. Using an indicator of the position of the political left on the issue of electoral system choice (based on data from the expert survey) and the effective electoral threshold as an indicator of the proportionality of the electoral system (Boix 1999; Cusack, Iversen, and Soskice 2007, 2010), we find a bivariate correlation of $r=0.83$ ($N=20$). Thus, we tend to find PR where the political left supported PR.

³ The difference increases when we remove respondents from the sample who consider themselves social scientists first and historians second (4.72 vs. 3.82, t-value of -2.42).

⁴ The difference between "pure historians" and historians with some social science background is significantly different from zero at the 5-percent level in case of the evaluation of social scientists' historical knowledge and the attention social science research pays to the historical dimension of political and social developments.

- Alesina, Alberto and Edward L. Glaeser. 2004. *Fighting Poverty in the US and Europe: A World of Difference*. New York: Oxford University Press.
- Blais, André, Agnieszka Dobrzynska, and Indridi H. Indridason. 2005. "To Adopt or Not to Adopt Proportional Representation: The Politics of Institutional Choice." *British Journal of Political Science* 35:1, 182–190.
- Boix, Carles. 1999. "Setting the Rules of the Game: The Choice of Electoral Systems in Advanced Democracies." *American Political Science Review* 93:3, 609–624.
- Boix, Carles. 2010. "Electoral Markets, Party Strategies, and Proportional Representation." *American Political Science Review* 104:2, 404–413.
- Burke, Peter. 2007. *History and Social Theory*. Cambridge: Polity Press.
- Cusack, Thomas, Torben Iversen and David Soskice. 2007. "Economic Interests and the Origins of Electoral Systems." *American Political Science Review* 101:3, 373–391.
- Cusack, Thomas, Torben Iversen, and David Soskice. 2010. "Coevolution of Capitalism and Political Representation: The Choice of Electoral Systems." *American Political Science Review* 104:2, 393–403.
- Cusack, Thomas, Torben Iversen, and David Soskice. 2010b. "Web Appendix For 'Coevolution of Capitalism and Political Representation: The Choice of Electoral Systems.'" http://www.people.fas.harvard.edu/~iversen/index_files/page0004.htm.
- Elman, Colin and Miriam Fendius Elman, eds. 2001. *Bridges and Boundaries: Historians, Political Scientists, and the Study of International Relations*. Cambridge MA: MIT Press.
- Kreuzer, Markus. 2010. "Historical Knowledge and Quantitative Analysis: The Case of the Origins of Proportional Representation." *American Political Science Review* 104:2, 369–392.
- Kreuzer, Markus. 2010b. "Response to Cusack, Iversen and Soskice's APSR Article and Web Appendix." <http://www41.homepage.villa-nova.edu/markus.kreuzer/>.
- Lijphart, Arend. 1994. *Electoral Systems and Party Systems: A Study of Twenty-Seven Democracies, 1945–1990*. Oxford: Oxford University Press.
- Thies, Cameron G. 2002. "A Pragmatic Guide to Qualitative Historical Analysis in the Study of International Relations." *International Studies Perspectives* 3:4, 351–372.

*Regime-Based Analysis in
Comparative Welfare State Research:
Old and New Approaches*

Michael Shalev
Hebrew University of Jerusalem, Israel
michael.shalev@gmail.com

For reasons that have been well-rehearsed elsewhere (e.g., Ragin 1987; Hall 2002; Shalev 2007), standard statistical methods are ill-suited to the tasks of characterizing and explaining cross-country diversity in systems of social protection.¹ This paper draws on several recent strands of methodological innovation in comparative welfare state research and related fields, with a view to encouraging more self-conscious and systematic use of Esping-Andersen's welfare state regimes as an

alternative epistemological framework for explanatory comparative research. I discuss diverse methodological tools and approaches which can help harness the power of a regime-based approach in empirical analysis. I introduce some tools belonging to the statistical family known as multivariate analysis (“methods that simultaneously analyze multiple measurements on each individual or object under investigation”; Hair et al. 1995: 5). Some such tools (factor analysis and cluster analysis) are already staples of the field, but the paper also shows the potential usefulness of correspondence analysis, which has some similarities to factor analysis but rests on assumptions more congenial to comparativists. The biplot, a “perceptual map” that makes it possible to simultaneously visualize affinities between cases and their attributes, adds value to these techniques.

A “comparative regime analysis” (Ebbinghaus 2008) is one which self-consciously builds on the testable implications of insights offered by the literature on welfare state regimes. These insights may concern specific welfare state ideal-types or the regime system as a whole. As Esping-Andersen’s (1990, 1999) foundational works illustrate, such an analysis may be quantitative or qualitative, exploratory or confirmatory. In weighing the empirical evidence, it may rely on the reasoning, expertise, and interpretive powers of the researcher, or on formal researcher-independent rules. The two most prominent formal methodologies in comparative welfare state research today, the regression family and the QCA (Qualitative Comparative Analysis) family, can both be utilized in regime-based research, but with limitations and without fully exploiting their core competencies. Comparative (often also historical) case studies, and the multivariate analysis family of quantitative techniques, are more naturally suited to the regime approach.

Regime-Based Analysis

Welfare state regimes can provide invaluable shortcuts for practitioners of comparative research. First, as distillations of the empirical record, ideal-typical regimes obviate the need for researchers to build up a comprehensive knowledge base from scratch, country by country. Second, and relatedly, the affinity between cases and types has already been documented in the welfare regime literature, aiding qualitative researchers to set manageable limits to the number of cases in their analysis, by purposively selecting those that exemplify regime differences. Third, and most importantly, regimes come pre-packaged with theoretical insights. Some of these concern the operating principles or “inner logic” of each regime, the glue that holds together multiple facets of policy and political economy. Others identify forces that are believed to account for regime differences. Still other insights are claims about the nature of regimes themselves—notably the argument that broad regime differences tend to be durable. No matter in what guise the theoretical claims of the regime literature appear, the key common denominator is that they may serve as a source of testable implications. If outcomes fit expectations, not only is Esping-Andersen vindicated, but even more important for the scholar in the trenches, she now has an empirically plausible explanation of cross-country heterogeneity. Yet, in keeping with the

assumptions of ideal-typical models, no one-to-one fit is demanded or expected. On the contrary, a regime-based analysis also invites both analysis of variation within regimes, and a focus on ambiguous or hybrid cases with affinities to more than one regime.

These empirical and theoretical assets furnished by the welfare regime literature have nourished three different types of comparative welfare state research that will now be introduced by example: systematic but non-formal case comparisons, formal quantitative analysis, and non-formal quantitative analysis. The last of these will provide our point of entry to multivariate methods. The fourth logical alternative, formal methods of comparing cases (QCA), is mentioned in the concluding section.

Regimes in Comparative-Historical Analysis

In his study of the causes and consequences of variation in early retirement policies in Europe, the U.S., and Japan, Ebbinghaus (2006) offers an illuminating illustration of the power of a regime-based approach in a close comparative study of the evolution of a particular policy area. The author’s methodological self-awareness makes this a particularly valuable resource (see also Ebbinghaus 2008). While following in the footsteps of iconic practitioners of comparative-historical research like Moore and Skocpol, rather than inventing a typology from scratch, Ebbinghaus exploits the advantages of standing on the shoulders of existing scholars. Since the topic of early retirement is located at the intersection of welfare states with firms and industrial relations systems, he mobilizes the resources of regime typologies that have developed more or less independently in all three areas, using them to purposefully select a manageable sample of illustrative cases as well as to develop both ex ante hypotheses and ex post interpretive insights. In keeping with what Ragin (1987) famously described as casual complexity, Ebbinghaus finds that variation both between and within country clusters can only be explained configuratively—in this case, by interactions between the logics of national welfare states, varieties of capitalist production, and systems of social partnership.

Regimes in Regression Analysis

With the proliferation of multi-country surveys based on standardized questionnaires, many researchers have adopted multi-level (hierarchical) models for simultaneously analyzing individual and country-level data. In this way, the membership of countries in welfare regimes can be integrated into conventional regression analysis of individual-level variables. Studies of this type seek to evaluate the impact of social policy regimes on key variables in welfare state research, most notably raising the methodological bar in studies of the determinants of public opinion towards inequality and redistribution (compare Svallfors 1997 with Jaeger 2006, 2009). Often the aim is to establish the net importance of regime effects, controlling for the variation between countries in population composition captured by demographic variables at the individual level. Some researchers also control for rival country-level explanations of cross-national variation. Arts and Gelissen (2001) offer a

classic illustration of this approach. They show that even after taking account of a diverse range of individual differences, the way that people rank three different principles of social justice varies systematically across welfare regimes.

Using the same basic design, Chung and Muntaner (2007) and Eikemo et al. (2008) illustrate a mushrooming research effort underway outside of the field of social policy, which is aimed at grading the performance of different welfare regimes on health and mortality indicators. In this sort of approach, regimes serve as a rather blunt instrument unless researchers take steps to pin down causality, either by incorporating causal mechanisms on the right-hand side of the equation, or developing sufficiently subtle outcome measures on the left-hand side (cf. Hurrelmann et al. 2011). Another way of tapping the power of multilevel models is by investigating cross-level interactions, which ask how regime membership affects relationships between micro-level variables. For example, a recent study by Mandel (2010) uses standard wage equations at Level 1 to isolate the effect of gender on an individual's location in his or her national wage hierarchy. The second level of the model makes it possible to show how the net gender gap depends on the welfare regime in which an individual is located. Then, in an attempt to tease out the causal mechanisms underlying such regime effects, policy indicators replace regime categories at Level 2 of the model.

This is a powerful approach with considerable potential, but several cautions are in order. First, the statistical power of the multi-level design is questionable in macro-micro comparative research. The number of countries analyzed at Level 2 (21 in Mandel's case) is infinitesimal compared to the Level-1 samples furnished by survey programs like the ISSP or ESS. As Snijders has noted, "sample size at the highest level is the main limiting characteristic of the [multi-level] design," and this problem is especially acute when testing for cross-level interactions (Snijders 2005: 1570).

Two other issues that deserve attention concern the fit between the analytical procedures invited by a multi-level regression model and the premises of welfare regime theory. First, as configurations of multiple attributes, regimes rest on the assumption that the meaning of any one policy depends on the overall policy configuration in which it is embedded. For instance, the corrosive effects of extended paid leave for women following childbirth on their future employment and earnings are likely to be far more severe in conservative than social-democratic welfare states. The reason is that the former are far more likely to offer incentives for fathers as well as mothers to take paternal leave, and to make subsidized public childcare and other incentives available to mothers who return to the labor force after a period of withdrawal. Consequently, unpacking regimes into discrete policies and modeling these policies as explanatory variables could actually end up sabotaging researchers' ability to identify the causal mechanisms responsible for regime effects.²

Another potential tension between the regime approach and multi-level modeling concerns the manner in which countries are assigned to regimes. Allocating each country to only one regime in which it has full membership is at odds with the

analytical status of ideal-types as condensations and caricatures of observable phenomena. When no uncertainty is permitted regarding country assignments this results not only in measurement error, but also in researchers forfeiting the leverage built into mixed and ambiguous cases. In welfare regime theory, such cases embody the interplay of contradictory political forces (cf. Esping-Andersen and Korpi 1984). They are even more central to Varieties of Capitalism theory, which argues that weak institutional coherence impairs the economic performance of hybrid cases.

Regimes in Non-Formal Quantitative Analysis

Informed by these limitations of multi-level modeling, another comparative study of welfare state effects on the gender wage gap (Mandel and Shalev 2009) also relies on quantitative analysis of micro-data to assess the evidence for predictions based on prior theoretical and empirical knowledge of welfare state regimes—but without attempting to link cause and effect in a formal statistical model. As I suggested earlier, the potential power of such an approach depends on the researcher's ability to identify specific testable implications of regime ideal-types, and to develop equally targeted outcome measures, so that observed linkages between hypotheses and outcomes can be plausibly interpreted as causal effects.

Mandel and Shalev's core argument is that previous research on gender wage gaps confounded the effects of class and gender inequality. They undertake to differentiate between these effects, at the hypothesis-building stage by exploiting the rich prior literature on social policy and family policy and their effects, and at the hypothesis-testing stage by decomposing the gender gap into class-inequality and gender-inequality components. Drawing primarily on Esping-Andersen's work, the authors lay the ground for their empirical analysis by demonstrating that three key indicators of the role of the state in class and gender inequality appear to confirm the existence of distinct clusters of countries (with the exception of one or two mixed cases). Finally, heuristic methods (scatterplots and a simple simulation) are employed to demonstrate the interplay between regime membership and the decomposed indicators of class and gender inequalities. Although the authors offer their own account of the consistency of this evidence with their theoretical expectations, and carry out a formal test of one potentially confounding influence (women's selectivity into employment), it is ultimately up to the reader to make his or her own judgment regarding the convincingness of the exercise.

Multivariate Statistical Methods

As made explicit in the last of the three studies reviewed above, it is a precondition for regime-based analysis that researchers have good reason to believe that policies actually do congeal into the distinct bundles posited by welfare regimes (and/or other regime types). These empirical configurations then play an essential role in determining the regime membership of country cases. In the wake of both the extensive reliance on Esping-Andersen's regimes in welfare state research and the controversy they have generated, these tasks

have turned into a lively research industry with a life of its own. I make no attempt to review this literature here.³ Instead I will focus on studies that have employed several related methods—factor analysis (FA), principal component analysis (PCA), and multiple correspondence analysis (MCA).

Factor Analysis and Multiple Correspondence Analysis

In earlier work (Shalev 1996, 2007), I argued that factor analysis is particularly appropriate for accomplishing the twin tasks of identifying welfare state regimes and assigning countries to them, because it directly addresses Esping-Andersen's core theoretical claim that two dimensions of variation drive observed differences in national welfare states. These two dimensions represent the extent to which policy reflects tension between socialism and economic liberalism, and the particular brand of conservatism that developed in Continental Europe. Practically speaking, I suggested that far from being daunted by the perennial problem of "too many variables," data reduction techniques like FA put them to good use.

A landmark study by Hall and Gingerich (2009) followed a similar approach. Theoretically, they posited the existence of two latent continua representing the two spheres of the political economy (corporate governance and labor relations) which Hall and Soskice (2001) maintain are the sites of the most important institutional complementarities. Graphically arraying the two sets of factor scores, as I did, led them to argue that the twofold typology is broadly justified. In a further step, Hall and Gingerich utilized variable-oriented methods to garner empirical support for a number of testable implications of these twin complementarities.

Other studies following the same strategy have utilized PCA (e.g., Tepe, Gottschall, and Kittel (2010), which tends to yield similar results to FA although it rests on somewhat different theoretical foundations and uses different measures of association. Analyses based on these methods suffer from two significant limitations. First, users of FA typically gear their use of it to the goal of identifying clearly separated dimensions (e.g., by "rotating" factors to enhance "orthogonality," absence of correlation, between them). This requires making strong theoretical assumptions about regime separation, similar to my reading of Esping-Andersen as claiming that conservatism is equally alien to both the liberal and social-democratic regimes. In fact, however, we know for instance that for different reasons conservatism resembles liberalism in its distaste for public service provision, and social democracy in its aversion to inequality. Secondly, the statistical assumptions made by FA and PCA are highly constraining. Indicators must be measured on a continuous scale and are assumed to be normally distributed, and the relationships between them are expected to be linear. Moreover, to obtain reasonable levels of uncertainty, the number of cases ought to vastly exceed the number of variables, bringing the "medium-N" problem painfully back in.

Nonlinear counterparts to PCA, of which the best-known is Multiple Correspondence Analysis, make it possible to escape these constraints.⁴ MCA generates a "map" based on the correspondence between the rows and columns of a cross-

tabulation.⁵ While designed, like PCA, to reduce variance in the data to two or more underlying dimensions, its internals are different. First, as a nonparametric technique, MCA makes no assumptions about the distribution of indicators or their associations. Second, indicators are categorical rather than continuous. If our original attribute measures are continuous variables they will be segmented into categories, making it easier to identify discontinuous effects. Technically, instead of a correlation matrix, which is the input to PCA, the raw material of MCA is a matrix of chi-squared distances between the categories of all attribute variables. MCA is known to many social scientists because of the "map" presented in Bourdieu's 1976 classic *La Distinction* visualizing empirical linkages found in survey data between individuals' tastes in food, music, etc. and their membership in specific social classes. The foundational text in English is Greenacre (1984).

The earliest application of this approach to regime-based welfare state research is a study by Wildeboer Schut and colleagues (2001) using nonlinear PCA, a technique similar (and sometimes identical) to MCA. Using a database comprising 58 different indicators for 11 countries, the study showed that the Netherlands was a hybrid case which uniquely failed to coalesce into any of the standard three welfare regimes because it shares features with two of them. More recently, an ongoing study by Ferragina, Seeleib-Kaiser and Tomlinson (2011) uses MCA to test the applicability of Esping-Andersen's regime clusters over time and across different policy areas.⁶

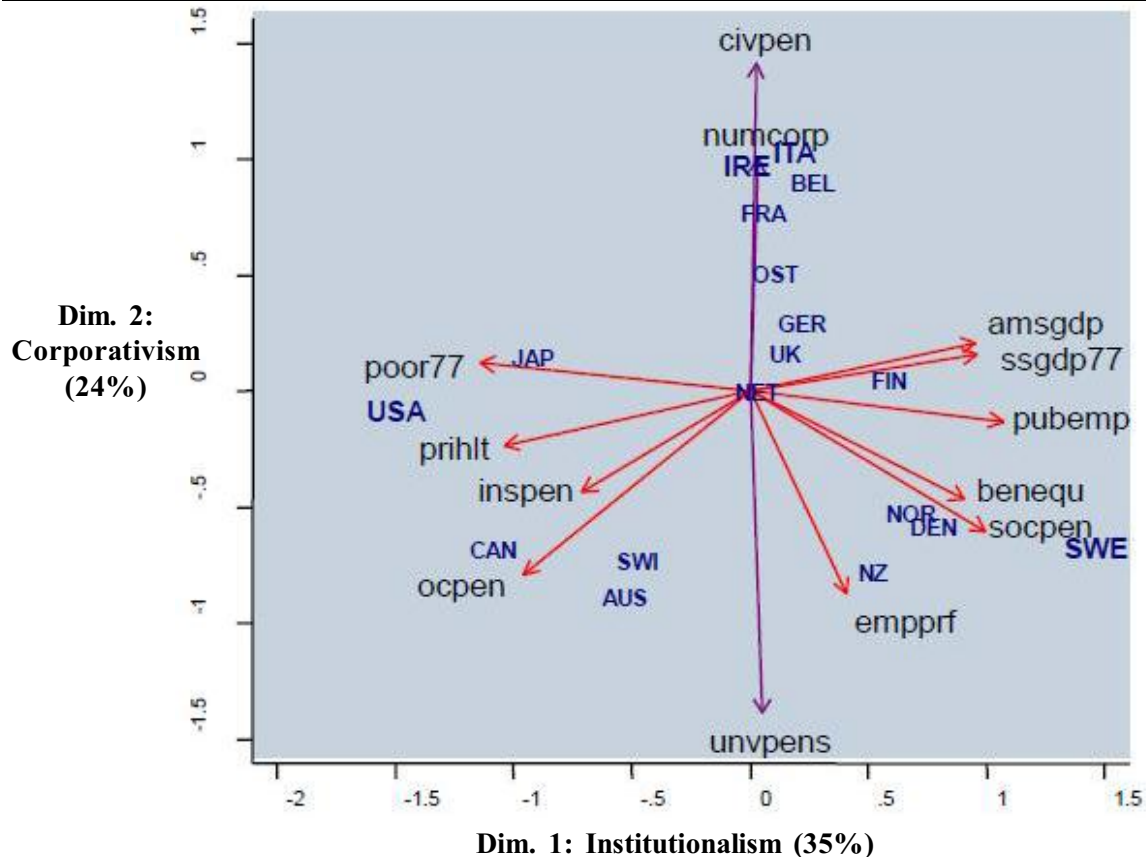
As in the case of my reanalysis of Esping-Andersen's data using FA, nonlinear PCA and MCA provide measures (comparable to factor loadings and scores) that can be presented in two-dimensional charts, showing affinities between countries or indicators. Relying solely on the country view, Ferragina et al. test the stability of welfare regimes by comparing the clusters which emerge when selected regime indicators are measured at three different points in time. This approach has the benefit of generating strikingly clear findings. However, it misses the opportunity to exploit a method of visualization, known as the biplot, which presents both case and attribute affinities in a joint two-dimensional space.

The Biplot

The biplot is a tool which has been reinvented several times. The "bi" in biplot refers to the juxtaposition of proximities between indicators and attributes. Biplots can be constructed for all variants of multivariate analysis, irrespective of differences in scale of measurement, distributional assumptions, and method of calculating proximities (Gower, Gardner-Lubbe, and Le Roux 2011; Greenacre 2010). The term biplot was originally used by Gabriel (1971) in the context of visualizing the results of PCA. In this context arrows are used to "project" the direction and strength of correlations between indicators, and cases are superimposed in locations that maximally reproduce their profiles on the set of measured indicators. In MCA, where indicators are measured categorically rather than continuously, each category is displayed as a separate point on the so-called "correspondence map."

Because successful biplots are uniquely able to portray

Figure 1: PCA Biplot of Esping-Andersen (1990) Dataset



affinities between countries and indicators, they can be distinctly useful in regime-based research for clarifying which indicators and countries are most germane to clusters, and assisting researchers to both identify and interpret ambiguous cases. For illustrative purposes, I have carried out both a PCA and an MCA of the same 13 indicators for 18 countries analyzed earlier using factor analysis. Due to space limitations, it is not possible to present the MCA chart here. However, a web appendix is provided⁷ which includes the dataset, the separate plots of factor loadings and scores published in Shalev (2007), and both the PCA and MCA biplots to which I shall now refer.

Comparing the two biplots, it is striking that even though MCA does not make the same stringent assumptions as PCA concerning the measurement and distribution of indicators or the form of their relationship, and does not require any excess of cases to indicators, overall the results obtained when applying both methods to the original *Three Worlds* dataset are very similar. The same three country clusters are found, according well with Esping-Andersen's own assignments. The same three ambiguous cases are evident (UK, Netherlands, and Finland). In addition, Australia is aligned with the Liberal countries and New Zealand is close to the Social Democracies—probably due to the absence of indicators of what Castles (1985) called “social protection by other means” (home ownership and protectionism). The only noteworthy difference between the two biplots in the location of countries concerns the Conservative cluster, which is more cohesive in the PCA version. In addition, in the PCA results Belgium is located within

the core group while Germany drifts towards the ambiguous cases, but the reverse occurs using MCA.

To interpret the vectors (arrows) in the PCA biplot, note that correlations between indicators are represented by the angles between the vectors: Right angles indicate lack of correlation, closely aligned vectors are highly positively correlated, and widely separated vectors (at the maximum, 180 degrees) are strongly negatively correlated. In addition, the longer the vector, the more reliable its correlation. Not surprisingly, the strongest correlation in this biplot is between two long and tightly connected vectors, for *civpen* (the share of pension spending on civil servants) and *unvpens* (the pension universalism), $r = -.81$. Note also that for a given case (country), its score on any indicator is ideally perpendicular to that indicator's vector. For instance, the three highest values of *numcorp* (number of separate pension plans) are for Italy, France, and Japan.

Once one becomes accustomed to its geometry, the PCA biplot is easier to interpret because it is more parsimonious. However, only MCA can identify potential nonlinearities. To capitalize on this, all of Esping-Andersen's original measures were split into three equal categories (tertiles). Substantively, the greater detail furnished by the MCA yielded two specific benefits. First, having an intermediate category on the map makes for an intuitive way of distinguishing between two types of regime distinctions: when attributes separate one regime from both of the others (e.g., the magnitude of civil service pensions), and when attributes carry a different weight in each

regime (e.g., “poor relief”). Second, and more important, while the PCA biplot offers scant indications of what accounts for the neutral positioning of the ambiguous cases, the MCA version points to some distinctive features of two of them. For example, the UK is characterized by medium and the Netherlands by high levels of spending on social security and active manpower policies, and Britain also stands out for the limited role of private health plans. The odd placement of the Finnish case reflects the hybrid character of its welfare state at the time—markedly conservative in pension fragmentation but highly social-democratic in other respects.

The lesson is that it is not always possible to squeeze all of the relevant information into a two-dimensional chart, and it is always advisable to return to the original data in order to verify that the spatial location of indicators and cases has not been misrepresented due to conflicting demands. In my view, this combination of formal (rule-based) manipulations, the parsimony afforded by data reduction and visualization, and the ability and indeed necessity of weaving back and forth between summary representations and raw data offers an attractive balance between “black box” techniques and “getting down and dirty” with data.

Conclusion

This paper has sought to clarify both the shared and different features of different strategies belonging to a single family of regime-based approaches to comparative welfare state research. It has also demonstrated that there is a wide range of analytical techniques, many of which cross-cut traditional methodological divides, which can be mobilized in the service of regime-based comparisons. To some extent the choice of strategies and techniques is a matter of researcher expertise and preferences, but notice also that due to the tradeoffs that often characterize these choices, researchers can profit from being methodologically ecumenical.

Space limitations prevent me from properly addressing two lacunae. First, I have not attended here to the vibrant methodological toolkit developed by Charles Ragin to approximate qualitative, case-oriented methods of comparative analysis. These methods have traditionally been used to identify the same types of elective affinities summarized by welfare regimes—but directly from the data. However, Kvist (2007) has argued that FSA can be equally well applied as an aid to developing and testing ideal-types of policy configurations. This method has since been adopted in several studies that sought to establish whether productivist or “workfare” policies conform with or contradict Esping-Andersen’s typology (Vis 2010; Hudson and Kühner 2009). While Hudson and Kühner (2010) have recently suggested that multivariate methods do a far inferior job of this than Kvist’s variant of fs/QCA, I believe that the matter hinges more on different interpretations of what ideal-types are and how they are connected to comparative analysis in regime-based research.

A second and more direct challenge to my suggestion to harness multi-variate methods in the service of regime-oriented research is a recent methodological critique by Ahlquist and Breunig (2009) of work on Varieties of Capitalism and welfare

regimes, accusing researchers of “reifying” these typologies. Using model-based clustering, these authors are unable to replicate previous findings, which it is claimed rest on the use of inadequate techniques that are merely “exploratory” (Ahlquist and Breunig forthcoming: 25–26). This critique should be constructively turned around, to encourage researchers to substitute the probabilistic measures generated by the techniques favored by these critics for absolute assignments of countries to regimes. At the same time, in practical applications these methods are far more ambiguous and dependent on researcher judgment than Ahlquist and Breunig admit. While the aspiration of reaching more precise and less arbitrary truth claims is always welcome, and methodological creativity based on borrowing advances made in other fields clearly has the potential to contribute to that goal, we should not forget that social scientists are for the most part limited to the (honorably) task of generating *plausible accounts* of what goes on around us.

Notes

¹ The author is grateful to Patrick Emmenegger for helpful suggestions.

² In theory, interdependencies between different state interventions could be handled by including conditional relationships (interactions) in the model. Unfortunately, severe degrees-of-freedom constraints rule out this possibility.

³ For an up-to-date review of the methods used by previous researchers for validating and operationalizing the three worlds, see Ferragina, Seeleib-Kaiser, and Tomlinson (2011).

⁴ For useful integrative treatments of the linear and nonlinear variants of PCA, see Meulman et al. (2004).

⁵ The adjective “multiple” signals that more than two variables are included in the cross-tabulation.

⁶ For an insightful application of both PCA and MCA in relation to questions raised by the Varieties of Capitalism typology, see Tepe et al. (2010).

⁷ http://pluto.huji.ac.il/~mshalev/Shalev_web_appendix.pdf

References

- Ahlquist, John S. and Christian Breunig. 2009. “Country Clustering in Comparative Political Economy.” Max Planck Institute (Cologne) Discussion Paper 09/5.
- Ahlquist, John S. and Christian Breunig. Forthcoming. “Model-Based Clustering and Typologies in the Social Sciences.” *Political Analysis*.
- Arts, Wil and John Gelissen. 2001. “Welfare States, Solidarity, and Justice Principles: Does the Type Really Matter?” *Acta Sociologica* 44:4, 283–299.
- Castles, Frances. 1985. *The Working Class and Welfare: Reflections on the Political Development of the Welfare State in Australia and New Zealand, 1890–1980*. Wellington: Allen and Unwin.
- Chung, Haejoo and Carles Muntaner. 2007. “Welfare State Matters: A Typological Multilevel Analysis of Wealthy Countries.” *Health Policy* 80:2, 328–339.
- Ebbinghaus, Bernhard. 2006. *Reforming Early Retirement in Europe, Japan, and the USA*. New York: Oxford University Press.
- Ebbinghaus, Bernhard. 2008. “Comparative Regime Analysis: Early Exit from Work in Europe, Japan, and the USA.” In *Methods and Substance in Macromethodological Analysis*. Basingstoke, UK: Palgrave.

- Eikemo, Terje Andreas, Clare Bamba, Ken Judge, and Kristen Ringdal. 2008. "Welfare State Regimes and Differences in Self-Perceived Health in Europe: A Multilevel Analysis." *Social Science & Medicine* 66:11, 2281–2295.
- Esping-Andersen, Gøsta. 1990. *The Three Worlds of Welfare Capitalism*. Cambridge: Polity Press.
- Esping-Andersen, Gøsta. 1999. *Social Foundations of Postindustrial Economies*. Oxford: Oxford University Press.
- Esping-Andersen, Gøsta and Walter Korpi. 1984. "Social Policy as Class Politics in Post-War Capitalism: Scandinavia, Austria, and Germany." In *Order and Conflict in Contemporary Capitalism*. J.H. Goldthorpe, ed. Oxford: Oxford University Press.
- Ferragina, Emanuele, Martin Seeleib-Kaiser, and Mark Tomlinson. 2011. "Welfare Regime Theory: A House of Cards? Unemployment Protection and Family Policy at the Turn of the 21st Century." *Annual Conference of the American Sociological Association*. Las Vegas.
- Gabriel, K. R. 1971. "The Biplot Graphic Display of Matrices with Application to Principal Component Analysis." *Biometrika* 58:3, 453–467.
- Gower, John, Sugnet Gardner-Lubbe, and Niel Le Roux. 2011. *Understanding Biplots*. West Sussex, UK: John Wiley & Sons.
- Greenacre, Michael. 2010. *Biplots in Practice*. Bilbao: Fundacion BBVA.
- Greenacre, Michael J. 1984. *Theory and Applications of Correspondence analysis*. London; Orlando: Academic Press.
- Hair, Joseph F., Rolph E. Anderson, Ronald L. Tatham, and William C. Black. 1995. *Multivariate Data Analysis: With Readings*. Upper Saddle River, NJ: Prentice-Hall, Inc.
- Hall, Peter A. and Daniel W. Gingerich. 2009. "Varieties of Capitalism and Institutional Complementarities in the Political Economy: An Empirical Analysis." *British Journal of Political Science* 39, 449–482.
- Hall, Peter A. 2002. "Aligning Ontology and Methodology in Comparative Research." In *Comparative Historical Research in the Social Sciences*. James Mahoney and Dietrich Rueschemeyer, eds. New York: Cambridge University Press.
- Hall, Peter, and David Soskice, eds. 2001. *Varieties of Capitalism: The Institutional Foundations of Comparative Advantage*. Oxford: Oxford University Press.
- Hudson, John and Stefan Kühner. 2009. "Towards Productive Welfare? A Comparative Analysis of 23 OECD Countries." *Journal of European Social Policy* 19:1, 34–46.
- Hudson, John and Stefan Kühner. 2010. "Beyond the Dependent Variable Problem: The Methodological Challenges of Capturing Productive and Protective Dimensions of Social Policy." *Social Policy and Society* 9:2, 167–179.
- Hurrelmann, Klaus, Katharina Rathmann, and Matthias Richter. 2011. "Health Inequalities and Welfare State Regimes. A Research Note." *Journal of Public Health* 19:1, 3–13.
- Jaeger, Mads Maier. 2009. "United But Divided: Welfare Regimes and the Level and Variance in Public Support for Redistribution." *European Sociological Review* 25:6, 723–737.
- Jaeger, Mads Meier. 2006. "Welfare Regimes and Attitudes Towards Redistribution: The Regime Hypothesis Revisited." *European Sociological Review* 22:2, 157–170.
- Kvist, Jon. 2007. "Fuzzy Set Ideal Type Analysis." *Journal of Business Research* 60:5, 474–481.
- Mandel, Hadas. 2010. "Winners and Losers: The Consequences of Welfare State Policies for Gender Wage Inequality." *European Sociological Review*.
- Mandel, Hadas and Michael Shalev. 2009. "How Welfare States Shape the Gender Pay Gap: A Theoretical and Comparative Analysis." *Social Forces* 87:4, 1873–1911.
- Meulman, J.J., A.J. Van der Kooij, W.J. Heiser, and D. Kaplan. 2004. "Principal Components Analysis with Nonlinear Optimal Scaling Transformations for Ordinal and Nominal Data." *The Sage Handbook of Quantitative Methodology for the Social Sciences*. 49–72.
- Ragin, Charles C. 1987. *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*. Berkeley: University of California Press.
- Shalev, Michael. 2007. "Limits and Alternatives to Multiple Regression in Comparative Research." *Comparative Social Research* 24, 265–313.
- Shalev, Michael, ed. 1996. *The Privatization of Social Policy? Occupational Welfare and the Welfare State in America, Scandinavia, and Japan*. London: Macmillan.
- Snijders, Tom A. B. 2005. "Power and Sample Size in Multilevel Linear Models." In *Encyclopedia of Statistics in Behavioral Science*, Vol. 3, B. S. Everitt and D. C. Howell, eds. Chichester, UK: Wiley.
- Svallfors, Stefan. 1997. "Worlds of Welfare and Attitudes to Redistribution: A Comparison of Eight Western Nations." *European Sociological Review* 13:3, 283–304.
- Tepe, Markus, Karin Gottschall, and Bernhard Kittel. 2010. "A Structural Fit Between States and Markets? Public Administration Regimes and Market Economy Models in the OECD." *Socio-Economic Review* 8:4, 653.
- Vis, Barbara. 2010. *Politics of Risk-Taking: Welfare State Reform in Advanced Democracies*. Amsterdam: Amsterdam University Press.
- Wildeboer Schut, J. M., J. C. Vrooman, and Paul de Beer. 2001. *On Worlds of Welfare: Institutions and their Effects in Eleven Welfare States*. The Hague: Social and Cultural Planning Office.

Announcements

Letter to the Editor (September 2011)

We are long-term readers of *QMMR*, and we find it overall an outstanding collection of essays and other materials related to methodological and methods issues in the discipline. Because of our high expectations, we would like to respond to John Gerring's reading of comparative politics and the impact of interpretive work in it. He wrote, in passing, in the Spring 2011 issue (vol. 9/1):

I will not address qualitative work in the interpretive tradition, as this tradition—while it certainly has its own version of perfection—is not as influential at the present time. (p. 8, Italics added)

From where we stand, this is a narrow reading of Gerring's own subfield terrain. Although we do not think that he necessarily had to engage interpretivist thinking in his essay, and while we do not believe it was his intention to slight interpretivist research, we nevertheless would like to take this opportunity to point out the not-inconsiderable achievements of work done in an interpretive vein.

Consider, for instance, the work of these comparativists done "in the interpretive tradition" (listed in the order of their initial publication, based upon the first work in the case of multiple publications by a single author):

James C. Scott, *Weapons of the Weak: Everyday Forms of Peasant Resistance* (Yale University Press, 1985); *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed* (Yale University Press, 1997);

Benedict Anderson, *Imagined Communities: Reflections on the Origins and Spread of Nationalism*, rev. ed. (Verso, 1991);

Frederic C. Schaffer, *Democracy in Translation* (Cornell University Press, 1998);

Lisa Wedeen, *Ambiguities of Domination: Politics, Rhetoric and Symbols in Contemporary Syria* (University of Chicago Press, 1999); *Peripheral Visions: Publics, Power, and Performance in Yemen* (University of Chicago Press, 2008);

Lloyd Rudolph and Susanne Rudolph, *Reversing the Gaze: The Amar Singh Diary, A Colonial Subject's Narrative of Imperial India* (Oxford University Press, 2000); and *Postmodern Gandhi and Other Essays: Gandhi in the World and at Home* (University of Chicago Press, 2006), among other works; and

Edward Schatz, ed., *Political Ethnography* (University of Chicago Press, 2009); **co-winner** of the QM Section's best book award last year.

Beyond comparative government, from the subfields of public policy and public administration/organizational studies and other branches of American politics, consider, to name but a few:

Robert Lane, *Political Ideology* (Free Press, 1967);

Graham T. Allison, *Essence of Decision* (Little, Brown, 1971);

Richard Fenno, *Home Style: House Members in Their Districts* (Little, Brown, 1978); **winner** of the 1979 Woodrow Wilson Foundation Award and the 1980 D.B. Hardeman prize, and re-issued in a "Longman Classics" Edition, 2009;

Michael Lipsky, *Street-level Bureaucracy: Dilemmas of the Individual in Public Services* (Russell Sage, 1980);

Jennifer Hochschild, *What's Fair? American Beliefs about Distributive Justice* (Harvard University Press, 1981);

Ann Chih Lin, *Reform in the Making: The Implementation of Social Policy in Prison* (Princeton University Press, 2000);

Joe Soss, *Unwanted Claims: The Politics of Participation in the U.S. Welfare System* (University of Michigan Press, 2000);

Steven Maynard-Moody and Michael Musheno, *Cops, Teachers, Counselors: Stories from the Front Lines of Public Service* (University of Michigan Press, 2003); **NSF-funded** research, **winner** of the 2005 Herbert Simon Book Award, APSA Public Administration Section, and the 2005 "Best Book" award, ASPA Section on Public Administration Research;

Katherine Cramer Walsh, *Talking about Politics* (University of Chicago Press, 2003); *Talking About Race: Community Dialogues and the Politics of Difference* (University of Chicago Press, 2007);

Dvora Yanow, *Constructing "Race" and "Ethnicity" in America: Category-making in Public Policy and Administration* (M. E. Sharpe, 2003); **winner** of the 2007 Herbert Simon Book Award, APSA Public Administration Section, and the first "Best Book" award, ASPA Section on Public Administration Research (2004);

Victoria Hattam, *In the Shadow of Race: Jews, Latinos, and Immigrant Politics in the United States* (University of Chi-

cago Press, 2007); **winner** of the 2007 Ralph Bunche Award from APSA; and

Anna Marie Smith, *Welfare Reform and Sexual Regulation* (Cambridge University Press, 2007); **winner** of the 2008 Victoria Schuck Book Award, APSA Women and Politics Section.

Then there are older works, classics, which are still read and which shaped the development of the field of study:

Philip Selznick, *TVA and the Grassroots* (University of California Press, 1949), reissued in the Classics of the Social Sciences series, 2011; *Leadership in Administration* (Harper & Row, 1957);

Peter Blau, *Dynamics of Bureaucracy* (University of Chicago Press, 1953);

Herbert Kaufman, *The Forest Ranger: A Study in Administrative Behavior* (Resources for the Future by Johns Hopkins Press, 1960); and

Michel Crozier, *The Bureaucratic Phenomenon* (University of Chicago Press, 1964).

And in IR, we offer a few of many we might have mentioned, to round out the picture:

Iver Neumann, *Uses of the Other: "The East" in European Identity Formation* (University of Minnesota Press, 1998);

Cecelia Lynch, *Beyond Appeasement: Interpreting Interwar Peace Movements in World Politics* (Cornell University Press, 1999); **winner**, Edgar J. Furniss Prize, Mershon Center on International Security, and **co-winner**, Myrna Bernath Prize, SHAFR;

Jutta Weldes, *Constructing National Interests: The United States and the Cuban Missile Crisis* (University of Minnesota Press, 1999).

Michael Barnett, *Eyewitness to Genocide* (Cornell University Press, 2003);

Ido Oren, *Our Enemies and US: America's Rivalries and the Making of Political Science* (Cornell University Press, 2003); and

Michael Loriaux, *European Union and the Deconstruction of the Rhineland Frontier* (Cambridge University Press, 2008).

We note that this list is by no means exhaustive, and we apologize to those left off it. With respect to the awards noted, we have identified those we knew of or could discover quickly, but we have not had the time to research this fully and may have missed some.

As our list makes clear, "qualitative work in the interpretive tradition" has, indeed, achieved prominence in comparative politics and other fields within the discipline. It is a vibrant and well-established tradition of social-scientific work, albeit one that develops different answers to the questions Gerring is posing in his essay. If he means that the interpretive tradition has relatively less influence than others within the political science discipline, we wholeheartedly agree: Neopositivism still rules the roost. But by framing his assessment of the interpretive tradition as he did, he missed a critical opportunity to educate readers of the newsletter and to open the door for further intellectually fruitful discussions about methods and

methodology.

Sincerely yours,

Kevin Bruyneel, Lee Ann Fujii, Patrick Thaddeus Jackson, Xymena Kurowska, Cecelia Lynch, Ido Oren, Timothy Pachirat, Edward Schatz, Ronald Schmidt, Sr., Peregrine Schwartz-Shea, Joe Soss, Dorian Warren, Cai Wilkinson, Dvora Yanow, and Cyrus Ernesto Zirakzadeh

**Workshop – Process Tracing:
Philosophy, Theory and Practice
ECPR Joint Workshops, Antwerp (April 10–15, 2012)**

The aim of the workshop is to contribute to the expanding methodological literature on qualitative case study methods. Our focus is on process tracing and in-depth case study methods more generally, as well as the role that they can play in multi-method research designs. Witnessing an increasing use of process tracing in empirical research and an increasing body of methodological literature on process tracing and causal mechanisms, we see the need to draw the literature together. At the same time, we need to develop further our knowledge on key aspects of process tracing, such as how Bayesian updating can be used as an inferential logic, and when and how process tracing studies can be nested in multi-method research designs.

We especially encourage papers on three different themes (the sub-themes are examples and not exhaustive):

(1) Challenges relating to the conceptualization and operationalization of causal mechanisms: What are causal mechanisms and how are they related to causal effects? How can they be conceptualized and operationalized in a manner that permits them to be studied in practice? What are the implications for process tracing that follow from specific definitions of causal mechanisms?

(2) Practice and methodological challenges of process tracing: Practical and methodological investigations of the evidential weight of different types of sources, comparing primary sources (archives) with forms of secondary sources (newspaper accounts, participant interviews) for their evidential weight. How can we assess empirical material using case-specific knowledge in a transparent manner that enables replication? What measures for the uncertainty of causal inferences are available in process tracing? What methodological implications does the Bayesian logic of inference have for process tracing?

(3) Using process tracing in combination with other methods in multi-method research: What is the epistemological common ground for the combination of large-N methods aiming for causal effects and process tracing searching for causal mechanisms? How can process tracing be used for the improvement of concepts, indicators, and the quality of measures in multi-method designs? Reaching beyond regression analysis, how can process tracing be combined with other techniques like social network analysis, Bayesian modeling, Qualitative Comparative Analysis (QCA), or cluster analysis? What are appropriate case selection strategies for process tracing in multi-method designs?

Qualitative & Multi-Method Research, Fall 2011

We encourage a broad range of participants, ranging from scholars working on developing methods to established scholars and Ph.D. candidates that are using process tracing methods as a part of their substantive research. For more detailed information, please check the website of the ECPR (www.ecprnet.eu) or get in contact with one of the workshop directors.

Co-directors: Derek Beach, University of Aarhus, Denmark, derek@ps.au.dk, and Ingo Rohlfing, University of Cologne, Germany, rohlfing@wiso.uni-koeln.de.

The David Collier Mid-Career Achievement Award

Recipient: Daniel Carpenter, Harvard University.

Committee: Colin Elman, Syracuse University; James Mahoney, Northwestern University, and Lisa Wedeen, University of Chicago.

The Award honors David Collier's contributions—through his research, graduate teaching, and institution-building—as a founder of the qualitative and multi-method research movement in contemporary political science. The award is presented annually to a mid-career political scientist to recognize distinction in methodological publications, innovative application of qualitative and multi-method approaches in substantive research, and/or institutional contributions to this area of methodology. To be eligible for the mid-career award, nominees must have defended their dissertation within 15 years of the beginning of the year in which the award is presented. For the 2011 award, nominees must have defended their dissertation in or later than 1996.

The 2011 David Collier Mid-Career Achievement Award was awarded to Daniel Carpenter of Harvard University. Carpenter has written on several methodological problems and techniques, including modeling learning and diffusion, learning from rare events, and stochastic prediction and estimation of nonlinear political durations. He has also served in various official capacities, including co-editing *Studies in American Political Development*. The Award Committee was, however, especially impressed by Carpenter's substantive research. His use of original archival data and interviews, coupled with statistical techniques and modeling, place him at the forefront of applied multi-method research. This track record has been established through a prodigious publication record, including dozens of articles and book chapters, and two major volumes: *The Forging of Bureaucratic Autonomy: Networks, Reputations and Policy Innovation in Executive Agencies, 1862–1928* (Princeton: Princeton University Press, 2001), and *Reputation and Power: Organizational Image and Pharmaceutical Regulation at the FDA* (Princeton: Princeton University Press, 2010).

The David Collier Mid-Career Achievement Award has been established by the Consortium for Qualitative Research Methods (the co-host of the Institute for Qualitative and Multi-Method Research). By agreement with CQRM, the award is managed by the APSA Organized Section for Qualitative and Multi-Method Research.

Giovanni Sartori Award for the Best Book Published in 2010 Developing or Applying Qualitative Methods

Recipient: Lauren M. MacLean, *Informal Institutions and Citizenship in Rural Africa: Risk and Reciprocity in Ghana and Côte d'Ivoire* (Cambridge University Press, 2010).

Committee: Daniel Brinks, University of Texas-Austin (chair); Heins Goemans, University of Rochester; and Adria Lawrence, Yale University.

Lauren MacLean marshals an impressive array of data, the product of extensive fieldwork in difficult conditions, in support of a complex argument that ties models of colonial and post-colonial state intervention to local informal institutions of reciprocity and exchange. These local processes are then connected back up again to national level politics and notions of citizenship. She finds that social reciprocity in Côte d'Ivoire is deep but tightly focused on the nuclear family, and argues counterintuitively that a more active state failed to foster broad patterns of reciprocity beyond the immediate family. On the other hand, what we might call wide but thin patterns of social reciprocity in Guinea are traceable to a less present state in the colonial and post-colonial era. These patterns are, in turn, associated with different conceptions of citizenship—communal vs. individual/entitlement-oriented, respectively—in each of these communities. These notions of citizenship aggregate to very different patterns of identity politics, which she later (more or less speculatively) connects to differing political outcomes at the national level. The argument, in short, connects national-level state variables to local-level micro-institutional outcomes, and these local outcomes to the national level again.

Her book stands out for its exemplary attention to methodological issues. In order to tease out the impact of colonial and post-colonial variables, she engages in very conscious and theoretically informed case selection. Her study relies on a matched-case examination of four villages in two countries, all inhabited by an ethnic group that was separated by colonially drawn national boundaries. Though these villages resembled one another closely during the precolonial period, they eventually produced different practices of reciprocity, making these communities quite different today.

MacLean uses a broad variety of methods to examine the different trajectories of these villages, blending macro-historical work on the state with ethnography, surveys, and focus groups on the local communities, using the findings of one to inform the others. Intensive field research provides a rare glimpse into local-level politics and practices at the village level. Archival material and oral histories supplement the field research. Moreover, she crosses disciplinary boundaries in search of the best methods for her research questions—for example, to anthropology, for best practices in ethnographic research. On several occasions she uses narrative to illuminate her research puzzle and argument. This combination of methods contributes nicely to supporting an argument that involves different units of analysis at different stages of the research.

What especially caught the attention of the committee, how-

ever, was the close attention to defining concepts and to the local boundedness of some of the concepts that are at the center of her research. In keeping with Sartori's insights on concept formation, she carefully defines some notoriously difficult concepts—e.g., informal institutions, norms—and self-consciously avoids concepts she finds too problematic for her purposes—e.g., social capital. In her discussion of the family, for instance, she acknowledges that notions of “nuclear” vs. “extended” family are neither indigenous to the area she studies nor uniformly understood across her villages (p. 48, 71). To give these concepts content, then, she relies on actual empirical investigation rather than assuming the same words convey the same meaning across geographic and cultural boundaries.

In summary, MacLean competently deploys a variety of methods in the service of an interesting and provocative research question, with close attention to concepts, in the best tradition of Giovanni Sartori's methodological advice. Her book well deserves the prize named in his honor.

Alexander George Award for the Best Article or Book Chapter Published in 2010 Developing or Applying Qualitative Methods

Recipient: Melani Cammett and Sukriti Issar, “Bricks and Mortar Clientelism: Sectarianism and the Logics of Welfare Allocation in Lebanon.” *World Politics* 62:3 (July 2010), 381–421.

Committee: Henry Hale, George Washington University (chair); Timothy Pachirat, New School for Social Research; and Joe Soss, University of Minnesota.

The committee is happy to present the 2011 Alexander L. George Award for the best article or book chapter developing and/or using qualitative research methods published in 2010 to “Bricks and Mortar Clientelism: Sectarianism and the Logics of Welfare Allocation in Lebanon,” by Melani Cammett and Sukriti Issar. “Impressive” was a word that repeatedly came up in our discussions of this piece. The authors address an important question in comparative politics: When do religious or ethnic organizations seek to mobilize support beyond their own groups? They do so by delving into an extraordinarily difficult and substantively important context, identifying and explaining the different strategies employed by the “Hezbollah” and “Future” movements in Lebanon. The result is an exemplar of how mixed methods can be convincingly deployed to address a substantive question of manifest real-world importance. The authors are able to pinpoint differences in how the two organizations distribute “bricks and mortar” patronage to in-groups and out-groups (in the form of clinics, hospitals, and private schools) by creating an original Geographic Information Systems (GIS) dataset and using statistical methods to relate the findings to local population characteristics. They then draw on over three hundred in-depth elite interviews (with interviewers carefully matched with interviewees) as well as archival and published sources to gain further insight into the motives behind these welfare provision strategies. The study is topped off with four case studies, including instances where

both movements have and have not set up welfare agencies. All this produces a nuanced and highly convincing argument that sheds new light on how sectarian groups mobilize supporters and when they will be more or less inclusive of out-group members. This article, in short, exemplifies the methodological excellence that our section is set up to encourage.

SAGE Prize for the Best Paper Presented at the 2010 Annual Meeting of the American Political Science Association Developing or Applying Qualitative Methods

Recipient: Adria Lawrence, Yale University. "Political Equality and Nationalist Opposition in the French Colonial Empire."

Committee: Stacie Goddard, Wellesley College (chair); Eleonora Pasotti, University of California-Santa Cruz; and Alison Post, University of California-Berkeley.

The winner of the 2011 Sage Award is Adria Lawrence, for her paper "Political Equality and Nationalist Opposition in the French Colonial Empire." Lawrence's paper asks why nationalist movements erupted in the 20th century French Empire, leading to the Empire's demise. She argues that it was not simply nationalism itself, but rather the denial of political equality to colonial subjects, that led to independence movements within France's colonies.

Lawrence's paper tackles a particularly bedeviling methodological issue, namely, how to rigorously evaluate counterfactuals in political science. In making her argument, Lawrence is suggesting that had France offered political rights in colonies such as Morocco and Algeria, then political mobilization within these colonies would have taken a far different path, a path that might not have necessarily ended in claims for independence from the Empire. But history, of course, does not provide us with alternative universes. French colonial subjects in Morocco and Algeria were never granted political equality and thus, as Lawrence notes, we have no straightforward route to test her claims.

Lawrence's efforts to construct a suitable test of her theory are remarkable, and should serve as a model for any political scientist interested in making counterfactual claims. She provides two distinct tests of her theory in the paper. First, she examines variation in nationalist movements at the macro-level of the French Empire. France did not extend political rights to Algerians and Moroccans, but they did incorporate a few colonies, as well as extended citizenship to particular polities in the 1946 *Loi Lamine Gueye*. In this section she demonstrates a correlation between granting political equality on the one hand, and the absence of nationalist movements on the other. Second, she identifies two colonies—Senegal and Spanish Morocco—in which there was striking sub-national variation, with some communes receiving political rights, and others denied. As with her macro-analysis, she finds that nationalist movements only occurred when political equality was denied. Lawrence's counterfactual analysis is impressive, not only for establishing variation, but for its efforts in process-tracing as well. Throughout the paper, Lawrence is not content to identify the presence or absence of nationalist movements. Extensive archival work underpins her analysis of decision-making,

and her efforts to establish why some individuals chose to press for independence, while others remained silent.

Had Lawrence's paper ended there, it would have been an excellent analysis of a significant substantive question. But she recognizes that her theory could raise charges of endogeneity—perhaps the French only granted political equality in those areas where nationalist movements did not exist, in which case, political equality had no causal effect. To address this methodological problem, Lawrence uses the final section of her paper to meticulously trace why the French offered political equality to some of their colonial possessions and not others.

Lawrence's paper, in short, is exemplary, both in its awareness of the methodological hurdles to her analysis, and in the execution of her case studies. Lawrence's paper should set a new standard for those of us in the field engaged in historical, counterfactual analysis.

Call for Nominations: The APSR Qualitative Submissions Award

The APSA Organized Section for Qualitative and Multi-Method Research is pleased to announce the establishment of the Qualitative Submissions to APSR Award, for the best qualitative manuscript submitted to the *American Political Science Review* in the calendar year.

The award will be offered in 2011 through 2014, and the winner in each year will receive \$2,000. To be eligible:

- (1) the manuscript need only be submitted to (not necessarily published in) the journal;
- (2) the manuscript needs to have been submitted during the calendar year, with the date of submission determined by the acknowledgement email from the APSR;
- (3) both new and subsequent submissions (e.g., resulting from an invitation to submit *de novo* or to revise and resubmit) are eligible for the award, but only one version of the manuscript is eligible for the award in any one calendar year;
- (4) the manuscript submitted to the APSR must be (a) new research on qualitative methodology *per se*, i.e., a study that introduces specific methodological innovations or that synthesizes and integrates methodological ideas in a way that is in itself a methodological contribution; and/or (b) substantive work that is an exemplar for the application of qualitative methods, or of multi-methods with a substantial qualitative component.

Nominations for the 2011 award should be submitted by January 31, 2012, via email to Gary Goertz (ggoertz@email.arizona.edu) the President of the APSA Organized Section for Qualitative and Multi-Method Research. Nominations need to include a copy of the submitted version of the manuscript version, and a copy of the APSR's email acknowledging receipt. Cover letters are optional. Self-nominations are welcome.

Qualitative and Multi-Method Research

The George Washington University
Department of Political Science
Monroe Hall 440
2115 G Street, N.W.
Washington DC 20052
USA

Nonprofit Org. U.S. Postage PAID TUCSON, AZ Permit No. 271
--

Qualitative and Multi-Method Research (ISSN 2153-6767) is edited by Robert K. Adcock (tel: 510-301-4512, fax: 202-994-7743, email: adcockr@gwu.edu). The production editor is Joshua C. Yesnowitz (email: jcyesnow@bu.edu). Published with financial assistance from the Consortium for Qualitative Research Methods (CQRM). Opinions do not represent the official position of CQRM. After a one-year lag, past issues will be available to the general public online, free of charge, at <http://www.maxwell.syr.edu/moynihan/programs/cqrm/section.html>. Annual section dues are \$8.00. You may join the section online (<http://www.apsanet.org>) or by phone (202-483-2512). Changes of address take place automatically when members change their addresses with APSA. Please do not send change-of-address information to the newsletter.