

# Causal Process “Observation”: Oxymoron or Old Wine

Nathaniel Beck\*

Draft of December 23, 2006

## ABSTRACT

The issue of how qualitative and quantitative information can be use together is critical. Brady, Collier and Seawright have argued that “causal process observations” can be adjoined to “data set observations.” This implies that qualitative methods can be used to add information to problematic quantitative data sets. In a recent symposium in Political Analysis I argued that such qualitative information cannot be adjoined in any meaningful way to quantitative data sets. In that symposium the original authors offered several defenses. Here I consider those issues more generally. I argue that the “causal process observations” are useful either for the explanation of specific events, or to help in the theory building process (or in any other way that qualitative analysts find them useful); but, they are not capable of being “adjoined” to standard quantitative data.

---

\*Department of Politics; New York University; New York, NY 10003 USA; [nathaniel.beck@nyu.edu](mailto:nathaniel.beck@nyu.edu). Thanks, for various things, to Henry Brady, David Collier, Gary King, Bernard Manin, Adam Przeworski and Jas Sekhon. A previous version was presented at the 2006 Annual Meeting of the American Political Science Association, Philadelphia, Sept., 2006.

In an important recent book, Brady, Collier and Seawright ([Brady and Collier, 2004](#)) argue for a unified methodology for both qualitative and quantitative social science that respects both traditions. Obviously it is hard to disagree with such a laudable goal. However, in a symposium on the book ([Beck, 2006](#)), I argued that the key contribution of the book, the joining of “causal process observations” to “data set observations” is chimerical. In a rejoinder to my comment, [Brady, Collier and Seawright \(2006\)](#) argued, using examples from natural science, epidemiology and political science that CPOs can be combined with DSOs in ways that I argued were impossible. Few are interested in the specifics of our disagreement, but many are interested in the vital issues raised by Brady, Collier and Seawright.<sup>1</sup> I thus rejoin some of the critical issues raised by BCS in the spirit of helping to better understand the relationship of qualitative and quantitative analysis.

Amongst the various important arguments made by BCS, the most important and novel is that analysts can join together both qualitative (“causal process”) and quantitative (“data set”) observations. There are many other issues raised in BCS; I will leave it to others to conduct what they find to be the most relevant dialogues on the important issues of unifying the various research traditions in our discipline. Suffice it to say, since this piece is about disagreements, that there are many suggestions in BCS not discussed here that I would find it hard to disagree with.

“Causal process observations” (CPOs) are the bread and butter of the qualitative analyst; “data set observations” (DSOs) serve the same role for the quantitative analyst. BCS ([Brady and Collier, 2004](#), 227-8) define a CPO as “[a]n insight or piece of data that provides information about context, process or mechanism, and that contributes distinctive leverage in causal inference. A causal-process observation sometimes resembles a “smoking gun” that confirms a causal inference in qualitative research, and is frequently viewed as an indispensable supplement to correlation-based inference in quantitative research as well.” The tie of CPOs to the qualitative analyst’s standard method is strengthened by adding to the definition a reference to “process tracing.”

CPOs are distinguished from DSOs; the latter are the quantitative researcher’s typical measures on a set of variables for each “subject” or “case” in the study. Obviously quantitative analysts find DSOs to be of value, and qualitative analysts likewise find CPOs to be of value. There is no need to debate this here. Clearly if we do two separate analyses, and each sheds some light, then the two together must shed more light than either one alone. The only debate is whether the two types of “observations” can be meaningfully combined.<sup>2</sup> But I take it the latter claim is what is novel in BCS.

Let me stress what is not at stake in the current paper. Researchers use a combination

---

<sup>1</sup>I will refer both to the original edited volume, the specific contributions of Brady, Collier and Seawright, and their reply generically as BCS (other than for quotations).

<sup>2</sup>[Brady, Collier and Seawright \(2006](#), 326) are willing to drop the term “observation” from CPO, noting that “[g]iven this potential misunderstanding as to whether researchers directly observe causal-processes[, if some readers find it more helpful to think of this as ‘causal-process information,’ that is a useful alternative label. However, to reiterate, we deliberately called these pieces of data causal-process observations to emphasize that this kind of evidence merits the same level of analytic and methodological attention as data-set observations in quantitative research. They continue to use the terminology of “adjoining these alternative kind of data” ([Brady, Collier and Seawright, 2006](#), 368).

of quantitative and qualitative methods, in differing proportions. Some find one or the other type of method of more use, but few would rule out either method *a priori*. Thinking primarily of comparative politics (broadly defined, so that it includes any types of comparisons across political units, and so includes almost all international relations and a good deal of the area study of American politics), the standard dissertation or book length study has a mix of quantitative and qualitative chapters (with additional purely theoretical chapters). Each of these chapters provides insight into some issue. And these analyses can be used to inform the others in an iterative manner (whether formally, as in [Lieberman, 2005](#), or more informally). What is at stake here is whether the two types of “observations” can be “adjoined” in a single analysis.

There is also no doubt that qualitative information is of great value to purely quantitative researchers. At a minimum qualitative information is useful in developing some basic understanding of how the world works. We expect students of comparative politics know their cases. To take even the most straightforwardly quantitative subfields of political science, our best scholars of the legislative process<sup>3</sup> have spent a year as Congressional Fellows, and our survey researchers pretest their instruments using, amongst other things, focus groups. In my own work on the politics of monetary policy, which in the end used only DSOs, my first step was to talk with various people inside and outside the Federal Reserve. While this does not guarantee that one will not still say some foolish things, clearly it does help to guard against some stupidities. Similarly, any American who has grown up outside Chicago would probably realize, as did Brady, that it would be very hard for any Floridian living in the panhandle counties to fail to vote based on an early call of the Florida outcome if they had already voted. Martians could be very good at running regressions but they would need some field experience before we would believe their quantitative analyses of Earth politics.

Since it is hard for me to know exactly what a CPO is, I begin by looking at the examples presented by BCS; these are from astronomy and paleontology, epidemiology as well as political science. While epidemiology looks a lot like quantitative social science, astronomy and paleontology do not, and hence I deal with the latter examples very briefly. All the various astronomical and paleontologic observations discussed look, to me, like straightforward “hard data.” (I would not call them DSOs since in these fields there is no data “spreadsheet.”) The observation of moons of Jupiter or various chemicals in the earth’s crust hardly look like what comes from process tracing.

## 1. NATURAL SCIENCE AND EPIDEMIOLOGY EXAMPLES

In many areas of science, where stochastic variation is irrelevant, scientists proceed by saying “if theory A is correct we must observe X, and so if we do not observe X, theory A must be incorrect.” So errors in the positioning of planets allowed for rejection of Ptolemaic theory (though of course the history of science is a bit more complicated than this). In this elementary philosophy of science textbook view, astronomers would subject Copernican

---

<sup>3</sup>On a personal note, as someone who studied with both Fenno and Riker at Rochester, I think everyone would agree that it was the joining up of the two different sets of interests of these scholars that allowed Rochester to have the impact on the study of Congress that it had.

theory to a similar set of tests, looking for implications of the theory and then seeing if they are upheld. This overly simple but not completely incorrect view is how many of us learned about the Einsteinian theories overthrowing the Newtonian theory: the famous Michaelson-Morley experiment measuring the speed of light under various conditions, or observations made during total eclipses of the sun. We have similar stories about the implications of the structure of DNA for what characteristics various X-ray crystallographs should have. So if BCS are simply saying that we should test the implications of our theories by examining whether their consequences are consistent with what we observe, it would be hard for anyone with any positivist view of science to disagree. But what this has to do with “causal process observation (or information)” is beyond me.

Epidemiology is more similar to social science. It has weaker theory and empirical work must take into account randomness. So the two examples cited by BCS, Semmelweis’ work on puerperal fever and Snow’s on cholera, are more relevant to our own discipline. Since Snow’s work is more well known in our discipline, I only go into details on Snow here.<sup>4</sup>

Snow is known for his carefully analyzed quasi-experiments showing that cholera was a water borne illness, caused by drinking water that had been contaminated by infected sewage. Snow clearly did make theoretical as well as empirical progress, advancing the contagion theory of cholera to argue that it was spread by something ingested, most likely infected water. The leading theory of the time was that cholera was caused by “miasma” (“bad air”). Understanding how cholera was spread was critical, since the miasma theory argued that London should be cleaned of its filth, whereas the water borne theory argued that it should deal with its contaminated (Thames) water supply.

Snow’s belief that cholera was a water borne illness was based on careful observation, particularly that the disease first affected the alimentary canal and that its rapid spread made simple person to person transmission an unlikely route for spreading the infection. We can argue about whether observing that the initial symptoms of a disease are related to the digestive system bears any relationship to the qualitative analyst’s process tracing, but these “CPOs” were not part of the crucial empirical work that Snow did to show that cholera was a water borne disease.

Snow showed that cholera was a water borne illness in two famous quasi-experiments. In one, he showed that users of a water company that drew from the contaminated portion of the Thames River suffered from cholera at a much higher rate than users of another company that drew from upstream, even though users of the two companies were so mixed that many neighbors used different companies. In the second, he showed that users of the contaminated Broad Street pump were much more likely to get cholera than users of other water sources.<sup>5</sup>

---

<sup>4</sup>Semmelweis is most well known for using experimental evidence to show the correct cause of puerperal fever. Hempel (2006, 169–70) states “[o]ther people before Semmelweis ... had correctly hypothesized about the nature of puerperal fever, and how it was being spread, although their message, like Snow’s on cholera, was largely ignored. Semmelweis is famous amongst epidemiologists then, not so much for the originality of his findings, as for the statistical, scientific approach that he brought to investigating the problem.” There is nothing in Semmelweis’ methodology that would give any pause to the most purely quantitative experimental researcher.

<sup>5</sup>We need to be careful about CPOs here. Brady, Collier and Seawright (2006, 364) state that the cholera “epidemic came to a close shortly after Snow convinced the authorities to remove the pump handle, thereby

The dangers of making conclusions without evidence is shown by David Freedman's (1991, 299) who notes the very positive with a bit of the negative. "Snow did some brilliant detective work on nonexperimental data. What is impressive is not the statistical technique but the handling of scientific issues. He made steady progress from shrewd observation through case studies to analysis of ecological data. In the end, he found and analyzed a natural experiment. (Of course, he also made his share of mistakes. For example, based on rather flimsy analogies, he concluded that plague and yellow fever were also propagated through water....)" It was only later purely quantitative tests, including the famous mosquito experiments of Walter Reed, that showed how yellow fever was actually transmitted.

Snow's empirical work was both novel and brilliant, but shows "only" excellent quantitative analysis rather than any challenges to such analysis. For example, he could not observe who used the contaminated Broad Street pump, so he plotted cholera outbreaks versus distance to the pump. But rather than using Euclidian distance, he took the distance it would actually take to travel to the pump over available streets; he was thus a precursor of the modern spatial analyst. In his analysis of users of the two water companies, he had to do incredibly hard (and clever) work to figure out which company serviced a given house. "Armed with his list of addresses where deaths had occurred, he travelled to south London and, like a detective solving a crime, he began weeks of pounding the streets and knocking on doors, asking the same question – from which company did the household get its water?" (Hempel, 2006, 172)

Excellent analyst that he was, he tried to find as many empirical implications of his theory and then see whether the theory was borne out in those case. For example, he famously found some brewery workers near the pump who did not get cholera, since they had no need for any water, and he found some residents of an almshouse who lived under shabby conditions but had access to purer water and also did not get cholera. Testing multiple empirical implications is clearly good but hardly controversial.

Both Freedman and Hempel use the metaphor of the detective. But in both cases, they are referring to careful observation. Snow's "detective" work or "shoeleather" epidemiology consisted of the gathering of hard, but very standard, DSOs. While the work was hard, it consisted "simply" of gathering, for a large number of households in the cholera affected area, the presence of absence of cholera and the water source for the household (measured either by which water company was used or closeness to the Broad Street pump). The modern day social science analogue of Snow as detective is the household researcher, carefully gathering information on household behavior. Hard work, careful work, but no challenge to standard quantitative analysis and no challenge to making inferences based on (good) DSOs. With this, we can move on to the political science examples.

---

preventing people from using the contaminated water .... This sequence of specific, localized CPOs, most dramatically the consequences of removing the pump handle, were diagnostic markers that linked the cause (contaminated water) to the effect (cholera)." Why the confirmation of a specific empirical prediction is a CPO escapes me; nothing in this discussion would give pause to the most quantitative of analysts. Thus, it is only amusing to note that Freedman (1991, 295–6) gives a different account. "As the story goes, removing the handle stopped the epidemic and proved Snow's theory. In fact, he did get the handle removed and the epidemic did stop. However, as he demonstrated with some clarity, the epidemic was stopping anyway, and he attached little weight to the episode."

## 2. POLITICAL SCIENCE EXAMPLES

BCS discuss four examples, three from their previous work and a new one, Lieberman's (2003) comparative study of the determinants of a nation's ability to raise taxes.<sup>6</sup> I have little to add on what I said previously on the Tannenwald (1999) and Stokes (2001) studies. Tannenwald, having four cases and no variation on the dependent variable, turns to documents to see the accounts that decision makers gave of why they failed to use nuclear weapons in a crisis. Reasonable people can differ on the utility of accounts given by decision-makers as to why they did what they did; sometimes they tell stories we like, and we are happy, and sometimes not. Tannenwald simply had no useful (quantitative) data and turned to a qualitative analysis, that is, she turned the question of why the US did not use nuclear weapons into what accounts did policy makers give as to why they did not use nuclear weapons. Both questions are interesting, but they are different.

Lieberman first ran a standard cross-country regression and found that tax revenues were well predicted by economic development. He then looked closely at two cases, South Africa and Brazil, with similar levels of economic development but very different tax revenues. He then hypothesized that elites are willing to pay taxes if they believe that they will receive a non-trivial portion of the subsequent government spending. Lieberman then tested this with another regression. The moves from case analyses to regression is not uncommon; as noted above, Martians do not typically run sensible regressions and students of Congress first observe it before they model it. But to say that Lieberman adjoined CPOs to DSOs seems to be stretching the idea of adjoining. While I am sure that there are some Martian political scientists out there mindlessly running regressions on comparative issues they do not understand, and while obviously Lieberman devotes relatively more effort to case study than do most quantitative political scientists, there is nothing in his work to challenge the standard ideas of inference (as in King, Keohane and Verba, 1994).

Stokes, like Lieberman, brings to bear both qualitative and quantitative evidence, and so again no challenge to current approaches. While we are all happy when political actors give an account of why they did what they did which is consistent with our theories, the value of these accounts is somewhat unclear. To take a simpler situation, imagine a study of why members of Congress facing serious scandals choose not to run for re-election. Imagine previous decades where there are very few such scandals, and the three or four scandalized all chose not to seek re-election. Would we be as happy with the conclusion that members of Congress facing a scandal chose to spend more time with their families as with the CPOs of Stokes (or Tannenwald)?

Finally, let us turn to Brady's reanalysis. Brady, Collier and Seawright (2006, 355–6) state

Brady seeks to make a causal inference about a singular outcome – i.e., whether the early media call of the election yielded a substantial vote loss by Bush in the Florida panhandle. Brady argues against the large-N, regression-based study carried out by Lott, which played a critical role in the national debate on the

---

<sup>6</sup>As in my earlier comments, my issues here are purely methodological and I have no interest in critiquing any of these works, which were chosen because BCS discussed them.

election, given Lott's conclusion that Bush did indeed suffer a major vote loss due to the early media call. Brady disputes Lott's findings, basing his analysis on CPOs rather than DSOs, as in Lott's analysis, utilizing a form of quasi-detective work focused on a sequence of steps that would have had to occur for the vote loss to be plausible. Far from offering an informal analysis that impressionistically suggests the linkages between the putative cause (the early media call) and the supposed effect (lower turnout), Brady presents a carefully structured evaluation focused on a sequence of necessary conditions. These necessary conditions are not met, and the causal claim presented by the large-N study is effectively ruled out. Contra Beck, this is an instance of successfully adjoining CPOs and DSOs to address the same problem of inference.

As noted in my earlier piece, there is no doubt in my mind that Brady's analysis is superior to that of Lott. Lott simply regressed county turnout in Florida in four national elections on fixed county and year effects, and, based on that simple two way analysis of variance, found that turnout in the Panhandle counties was approximately ten thousand voters fewer than predicted by the simple analysis of variance. This is a totally atheoretical regression, of the type that our hypothetical Martian political scientist might choose to run.

Brady, instead, noted that the early call came only a few minutes before the polls closed, and that only people who had not already voted could have been dissuaded from voting by the early call. While this is a causal argument, it follows from general knowledge of American voting (that one cannot go back and unvote), rather than quasi-detective work. Brady then assumes that people vote more or less uniformly over the day and so only a relatively small number of people might have been dissuaded. Surely the study of voting by time of day is a subject amenable to standard quantitative analysis. Brady then goes on to note that (quantitative) studies of media effects show that only a few are persuaded, and so forth. Brady then puts these together in a sequence (for the early call to have an effect, one must have not voted, heard the call, been persuaded by it, and the like) and then shows that since all these things are necessary for the call to have an effect, the turnout decrease claimed by Lott was most likely much too high.

Brady also provides an alternative explanation for Lott's regression results (which only, of course, could show that turnout in the Panhandle counties was lower than would be expected based on 2000 overall turnout and specific county effects). This explanation is based on Democratic mobilization in south Florida, which raised turnout in those counties (which would then raise expected turnout in the Panhandle counties in Lott's naive model). Of course it also might have rained in the Panhandle, or there could have been a major traffic accident, or any number of other idiosyncratic factors leading to the Lott result. While obviously turnout in Florida in 2000 was a politically (and perhaps legally) important issue, our interest as political scientists should be the determinants of turnout (including factors like party mobilization). These can then be applied to specific cases should interest warrant.

What of the claim that the carefully structured sequence of necessary conditions makes the Brady study the successful adjoining of CPOs to DSOs? That is, what this process has to do with "smoking guns" or "process tracing" is highly unclear. To make the issue simpler, let us return to epidemiology. Imagine we are interested in how many lives will be saved by

a new screening process. So we do a standard experimental study, randomly giving some people the screening process, some a placebo process. Both the false positive rate (how many unnecessary operations will be done) and the difference between the true positive and false negative rates (how many lives will be saved) can be computed. We then compute how many deaths are caused by unnecessary operations, and perhaps do another quantitative study on how many people will choose to be screened. Thus, using purely quantitative tools, and only DSOs, we can compute how many lives the screening program will save (or cost!). If a CPO is simply the an understanding that in life many will not choose to be screened, or that false positives are costly, then of course we need CPOs. But this is not the usage of CPO as defined by BCS. So while Brady did an admirable study, I fail to see how this is an instance of successfully adjoining CPOs to DSOs. Being committed to DSOs and quantitative analysis does not mean that one also must be committed to being a Martian political scientist!

### 3. CONCLUSION

What then can we make of the general argument? [Brady, Collier and Seawright \(2006, 359–60\)](#) ask why “are CPOs inherently important?” and then answer

... that a major part of the answer is found in the empirical and theoretical limitations of DSOs. DSOs are especially useful for detecting probabilistic relationships when there are many observations of comparable units and when the relevant causal pathways and interactions are well understood or controlled through randomized experiments. Yet it is often hard, if not impossible, to increase the number of analytically relevant DSOs, and we seldom have adequate understanding or control of these causal pathways or interactions.

In many situations, only a few similar units or events can be studied. .... In these situations, adding cases may be impossible (or at least foolhardy) because new cases will differ fundamentally from the original universe of concern. Furthermore, in some situations, social scientists – and also natural scientists – wish to explain a singular event .... In these situations, large-N studies may not be relevant.

The second limitation is even more serious. Increasing the number of DSOs can provide additional leverage when the major inferential problem is lack of statistical power stemming from a weak probabilistic relationship between a putative cause and effect. However, it can merely add cost without increasing leverage if the fundamental difficulty is a research design that diverges from a true experiment. In that case, the most vexing problem is not lack of data. Rather it is the lack of an appropriate design and the failure of a proper statistical specification that can allow for valid inferences. Proper statistical specification depends on strong theory that can guide the analysis.

Thus, the value of the regression-based model of inference, which is central to mainstream quantitative methods, depends upon a large number of comparable DSOs, which are hard to obtain, and on strong statistical specification assump-

tions, which are difficult to satisfy. We argue that a different strategy involving CPOs is sometimes more productive and less costly. This approach is hardly a panacea, because it can still demand careful data collection and the strong assumptions needed to interpret the CPOs. Yet it does provide an alternative when DSOs and regression analysis appear imperfect or defective.

There are serious criticisms here of the “regression” approach that I surely agree with. Sometimes we simply do not have enough comparable cases to do quantitative analysis. The solution is clearly not to add incomparable cases (though it may be to cleverly find a way to get more comparable cases). At that point the disappointing answer may be that we are not going to be able to do inferential research on this question. So we can turn the question around. Thus I think that Tannenwald’s question as to why the US did not use nuclear weapons after World War II may simply be impossible to answer; we can turn to a different question, that is, what accounts did policy-makers give of their decisions, but that is a different question.

Another issue relates to findings about general relationships (laws?) versus the explanation of specific events. In my previous critique I have argued that the former is what social scientists should be doing. But obviously there are different practices. As [Mahoney and Goertz \(2006\)](#) note, qualitative researchers are often interested in explaining single events while quantitative researchers are more interested in generalizable statements. While we surely can use general relationships to help explain specific events, the two enterprises are quite different. Thus standard experimental (and epidemiological) evidence can show that smoking leads to cancer, but it is very hard to know whether some particular person died of cancer because they smoked or because they worked in a shipyard (though our knowledge of medicine would allow us to dismiss “bad air” as a causal factor). While the latter question may be of interest in a courtroom, it is the former that should be of interest to us as scientists.

[Brady, Collier and Seawright \(2006, 360\)](#) use the metaphor of the detective.

At the simplest level, CPOs are diagnostic pieces of information that provide key insights in assessing explanations. A standard metaphor employed in discussing CPOs involves the parallel to criminal detective work. Detectives make their diagnoses on the basis of dogs that don’t bark (as in Sherlock Holmes’ famous “Silver Blaze” story), missing suicide notes, other clues that are or are not found at crime scenes, and stories that “just don’t add up.” ... this type of analysis goes beyond a simple model of “cause and effect” and recognizes that a causal process typically involves complex mechanisms, mediators, and markers that can provide alternative ways to test theories and to develop explanations. Paying attention to these mechanisms, mediators, and markers can reveal causal processes, and they are the foundation for CPOs.

This use of the detective metaphor is different from that of Freedman and Hempel.

I think that social scientists resemble criminologists, that is, seekers after general principles. These would be “dogs generally bark in the presence of strangers” or “suicides generally leave notes.” Establishing whether these principles are consistent with the behavior of dogs

or suicides is the stuff of standard quantitative analysis using only standard DSOs (though perhaps our criminologist first thought about dogs not barking by observing his own dog).

Detectives use these generalizations to help explain a specific case. Did Mr. X commit suicide? Unlikely, since no note was found. But specific cases admit to complicated causal stories. So perhaps Mr. X was a suicide but, alas, illiterate. Or maybe the dog that did not bark in the night was a Basenji. Explanations of individual events will also be complex and involve complicated causal stories. Why did the US invade Iraq? Clearly there are systematic explanations, but we can also find very specific stories related to highly non-systemic matters (say family relationships). My own take is our interest should be in systematic relationships, but it seems clear that whatever our interests, there is no adjoining of CPOs to DSOs to improve a defective research design.

Clearly there are regression models that range from theoretically unsatisfying to silly. While the assumption of linear additivity often provides a very convenient simplification, we should not always leap to this assumption. There are many standard statistical methods for going beyond linear additivity. So one can critique many quantitative studies without leaving the standard quantitative world. To give but one example, BCS make much of modeling a sequence of necessary conditions. But [Braumoeller \(2003\)](#) provides a standard maximum likelihood method for estimating models with necessary conditions. I agree with various authors in *Rethinking Social Inquiry* that much quantitative work is based on inadequate thinking; but, unlike BCS, I think that conventional quantitative tools, combined with good analytic thinking, are the solution here.

It is equally clear that more data do not solve critical problems of research design. The critical issue of endogeneity and selection bias are clearly not solved by more data. While in some areas of political science we can turn to experimental methods, this is almost impossible for the study of comparative politics. Sometimes we use clever quasi-experiments, akin to those of Snow ([Banerjee and Iyer, 2005](#)), but such opportunities do not always present themselves. Sometimes we can use good econometric methods to handle selection or endogeneity, though the assumptions behind such methods are very strong and often not met in the practice of comparative politics. But perhaps the conclusion should often be one of pessimism. While it is sad to end on the pessimistic note of [Przeworski \(Forthcoming\)](#) rather than the optimistic note of BCS, I fear the former is more accurate. Przeworski's answer to the question of what comparativists can do in the presence of endogeneity is "to try different assumptions and hope the results do not differ. .... If they do differ, all we can do is throw up our hands in the air. Where history was kind enough to have generated different causes under the same condition we will know more and know better. But history may deviously generate causes endogenously and this would make our task next to impossible." If we should be more optimistic than this, it is not because CPOs can remedy the defects of an inadequate research design.

## REFERENCES

- Banerjee, Abhijit and Lakshmi Iyer. 2005. "History, Institutions and Economic Performance: The Legacy of Colonial Land Tenure Systems in India." *American Economic Review*.
- Beck, Nathaniel. 2006. "Is Causal-Process Observation an Oxymoron?" *Political Analysis* 14:347–52.
- Brady, Henry E. and David C. Collier, eds. 2004. *Rethinking Social Inquiry: Diverse Tools, Shared Standards*. Lanham, Md.: Rowman and Littlefield.
- Brady, Henry E., David Collier and Jason Seawright. 2006. "Toward a Pluralistic Vision of Methodology." *Political Analysis* 14:353–68.
- Braumoeller, Bear F. 2003. "Causal Complexity and the Study of Politics." *Political Analysis* 11:209–33.
- Freedman, David A. 1991. "Statistical Models and Shoe Leather." *Sociological Methods* 21:291–313.
- Hempel, Sandra. 2006. *Medical Detective: John Snow and the Mystery of Cholera*. London: Granta.
- King, Gary M., Robert O. Keohane and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton: Princeton University Press.
- Lieberman, Evan S. 2003. *Race and Regionalism in the Politics of Taxation in Brazil and South Africa*. New York: Cambridge University Press.
- Lieberman, Evan S. 2005. "Nested analysis as mixed-method strategy for comparative research." *American Political Science Review* 99:435–52.
- Mahoney, James and Gary Goertz. 2006. "A Tale of Two Cultures: Contrasting Quantitative and Qualitative Research." *Political Analysis* 14:227–49.
- Przeworski, Adam. Forthcoming. "Is the Science of Comparative Politics Possible?" In *Handbook in Comparative Politics*, ed. Charles Boix and Susan C. Stokes. New York: Oxford.
- Stokes, Susan Carol. 2001. *Mandates and Democracy: Neoliberalism by Surprise in Latin America*. New York: Cambridge University Press.
- Tannenwald, Nina. 1999. "The Nuclear Taboo: The United States and the Normative Basis of Nuclear Non-Use." *International Organizations* 53:433–68.