

Is causal-process observation an oxymoron? a comment on Brady and Collier (eds.) “Rethinking Social Inquiry”

Nathaniel Beck*
Department of Politics
NYU
New York, NY USA
nathaniel.beck@nyu.edu

Draft of August 16, 2005

When I first read [King, Keohane and Verba \(1994\)](#), (“DSI” with authors “KKV”) I was excited by the general theme of what seemed to me the obvious, but usually unspoken, idea that all political scientists, both quantitative and qualitative, are scientists and governed by the same scientific standards. DSI, as Henry Brady cleverly notes, clearly was preaching to the qualitative researcher, and the subject of the sermon was that qualitative researchers should adopt many of the ideas standard in quantitative research. Thus, to abuse both the metaphor and KKV, one might compare DSI to the view of the Inquisition that we are all God’s children. But, having sat through many talks based on case studies, it is clear to me that getting everyone to think about the nature of the scientific enterprise we are all engaged in can only be a good thing.

[Brady and Collier \(2004\)](#) (“RSI” with authors “BC”) clearly takes no exception to this general position. But while DSI more or less argues that, *mutatis mutandis*, qualitative researchers should learn much from research design from quantitative ideas, RSI puts more emphasis on the *mutatis*, so that all God’s children can happily coexist, whether they worship one or many gods. To continue to abuse the metaphor, RSI clearly believes there are many ways to the father. My question for the current sermon is whether those many different paths are sufficiently similar. In particular, I am concerned about the role of “causal-process observations” in the scientific endeavor.¹ Obviously this is only one of many possible sermons; the other obvious sermon would be on “what the inquisitor can learn from the heathen?”²

*Thanks to Henry Brady, David Collier and Gary King, as well as my colleagues Bernard Manin and Adam Przeworski for interesting conversations on this topic.

¹This topic appears in most of the new work in RSI. For concreteness, I limit my discussion to the discussion in BC’s concluding chapter, found on pps. 252-66. The alert reader will note that the authors of this chapter are listed, in order, as David Collier, Henry Brady and Jason Seawright, but for simplicity I will continue to refer to the authors as BC.

²Thus RSI contains very important discussions on the limitations of even the best quantitative methods

“Causal-process observations” (CPOs) are, in my view, the key innovation in RSI. In particular, CPOs allow BC to argue that zero variance designs and selection bias are not as critical a problem as claimed by KKV. BC (pps. 277–8) 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.” BC distinguish CPOs from “data-set observations” (DSOs); the latter are the quantitative researcher’s typical measures on a set of variables for each “subject” in the study. BC criticize DSI for limiting discussion to DSOs. Clearly CPOs are the bread and butter of the qualitative analyst. What is novel in RSI is the linking of CPOs and DSOs and the argument that both types of information can, at least conceptually, be thought of as comprising the data matrix to be analyzed. Thus the insights of the qualitative analyst can be brought to bear on quantitative studies, and we can all live happily in a world of research which combines quantitative and qualitative data.³

So can we think of “adjoining” these CPOs to the DSO data matrix so as to circumvent problems clearly identified by KKV, problems which only arise because of limiting the data matrix to DSOs? To answer this for myself I looked at the three examples given by BC. I start with the most qualitative example and end with the most quantitative one. I stress that all discussion of these works is based on the discussion in BC, and so the authors of the original works may find that I am discussing work they have never seen. But it is BC’s methodology, not that of the original authors, that is at issue here.

The first example is of a purely qualitative assessment of why the US has not used nuclear weapons after World War II (Tannenwald, 1999), an in-depth analysis of three cases. No one would claim there is any variation in the dependent variable here. Clearly if we had some quantitative information on some independent variables (say party control of various branches of government), we could learn nothing by a regression of the dependent variable on those indicators. And clearly any variable that did not change over the course of the study could apparently explain US policy as well as the ideology variable proposed by Tannenwald. Since obviously observations on three cases (even if the dependent variable actually varied) would tell us little, Tannenwald focuses on documents about decisions to use nuclear weapons (process tracing). Clearly, as BC argue, this documentary evidence provides additional support for Tannenwald’s argument, but to what extent should we see this as additional “data,” albeit data of a different type?

Qualitative researchers would have no trouble with the type of evidence used by Tannenwald; the methodology she uses is almost prototypical. The issue is how to make sense of combining this qualitative evidence with the (in this case meager) quantitative evidence. On this issue BC are silent. I don’t think anyone would doubt that Tannenwald’s qualitative

which use observational data to make causal assessments, as well as mistakes quantitative researchers frequently make when they assume causal homogeneity. But the tolerance of the editors for multiple sermons is not high, and there are many other sermonizers in this issue.

³As is noted in various places throughout RSI, CPOs are related to information typically gained by such methods as “process tracing.” What is novel in RSI is the partial uniting of the qualitative and quantitative world, not any new tools unknown to either world.

evidence makes it somewhat more likely that her hypothesized mechanism is the correct one (or surely it does not make it less likely), but how do we go beyond that. As BC note, Tannenwald's evidence could be accounted for in other ways. They then conclude (p. 258) that "[h]owever, to the extent that researchers find alternative accounts such as strategic misrepresentation less plausible, Tannenwald's causal-process observations provide valuable support for her argument." To my mind this simply begs the question: good research design is about ruling out these alternative explanations, and BC provide no account of how this is to be done in the context of this study

The second study discussed is Stokes' (2001) study of the determinants of economic policy in Latin America from 1982–1995. According to BC (p. 256), she did a fairly typical comparative study of 38 cases, using both standard quantitative tools and more informal comparisons. She concluded that presidents chose neo-liberal policies because they believed such policies would solve fundamental problems. So far we are in the realm of standard DSOs and no interesting issues beyond those considered in KKV have yet arisen.

BC (p. 257) then note that Stokes "supplements the large-N analysis by examining a series of causal-process observations concerning three of the presidents....her inferential leverage derives from the direct observation of causal links." Thus she found that after his election, Fujimori encountered various leaders who exposed him to macroeconomic arguments which convinced him of the wisdom of adopting neoliberal policies. Clearly no one could be opposed to knowing one's cases, and clearly all social scientists do (or should) look carefully at a variety of cases before and after more systematic quantitative analysis. But what do BC mean by inferential leverage? Obviously we have all looked at some regressions and said they are nonsense, they simply miss some feature of the world; others seem more believable. Here one might trust Stokes since it appears that she does understand her cases. But can we infer beyond what we could infer from the 38 case regression? I do not see how.

The Stokes discussion raises two points. The first is whether our interest is in finding some general lawlike statements or in explaining a particular event. If the former, then knowing whom Fujimori talked to right after his election is not particularly interesting. But if the question is "why did Fujimori choose the neo-liberal policies?" then careful process tracing of the decision does seem relevant. I would argue that we are interested in the general lawlike statement, not the specifics of the Fujimori decision (other than it is covered by the lawlike statement).

The second relates to what it means to directly observe a causal link. We can ascertain whom Fujimori talked with, and we can certainly get accounts by those present as to why he did what he did. But should we take these at face value. While it is of interest to read the documents and hear the accounts of participants, it is difficult to know if these accounts are yielding the actual causal mechanism. It is not surprising that actors give accounts in terms of ideas, and that learning that policies are good policies is a common account of why policy-makers do what they do. As with Tannenwald, the issue is how we come to believe that these accounts are correct. Thus I do not know what it means to directly observe a causal process or how I know when a researcher observes such a process. Hence I think that Stokes work can only be evaluated by the standards set forth in KKV. I do not believe that KKV ever suggested it was wrong for analysts to understand their cases, or for them to have

done detailed studies to better understand those cases (or a subset of them). There is, of course, one more way to substantiate large-N studies, that is, by reference to theories that many people accept. I return to this issue after considering the third study discussed by BC.

This study is Brady's reanalysis (which appears as an appendix to RSI) of an unpublished study by John Lott on the impact of the early call of Florida for Al Gore in 2000. Lott argued that the call, ten minutes before the polls closed in the Florida panhandle, cost George Bush 10,000 votes. Lott came to this conclusion based on a time-series-cross-section regression of turnout in Florida counties over 4 elections. This regression was about as atheoretical as a regression could be, using as independent variables only county and year dummies. This regression was then used to conclude that Republican turnout was 10,000 voters under what would have been expected. (I rely here only on Brady's account, which is brief).

Brady argues that adding CPOs to the study would have led to a much more plausible estimate. The key feature here is that the call for Gore came only 10 minutes before the polls closed, and that prior standard quantitative studies indicates that people vote uniformly over the day; thus only a bit over 4200 voters in total (Republicans and Democrats) could have been dissuaded from voting by the early call. Brady combines this with prior empirical studies of the impact of media on turnout (including early calls) and concludes that Bush lost fewer than 100 votes due to the early call. Based on the evidence presented in the Appendix, it appears to me that Brady clearly bested Lott in this debate. But our interest here is in the role of CPOs, not the debate between Brady and Lott.

Brady brings to bear standard quantitative evidence such as survey evidence about when people vote over the course of a day and the amount of attention people pay to the media. In terms of his differences with Lott, the biggest impact is that the early call came only 10 minutes before the polls closed. We can either call it theory or common sense that the early call could not have affected anyone who had voted before the call. Thus Brady relies on standard quantitative evidence and I fail to see why CPOs are relevant here.

Brady concludes (p. 271) that "it would be hard to collect data that could rule out all the possible confounding effects. Consequently, rather than seeking additional data-set observations, in my judgement it would be more productive to do further in-depth analysis of causal-process observations drawn from these ten Florida panhandle counties, finding out what happened there, for example, by interviewing election officials and studying expert reports." As with Stokes, it clearly would be good for Brady to know as much about this case as possible. But election officials may have read Lott, and perhaps even have party affiliations. How would these CPOs provide any evidence to rule out anything? Better studies of turnout or media effects, of a standard quantitative kind, would provide more evidence. Why does Brady not suggest we do in-depth interviews of non-voters in the panhandle? Perhaps because we would not necessarily believe their own accounts of why they did not vote. As with the other articles, I remain in the dark as to what it means to observe a causal process.

Interestingly, where Lott appears to have gone wrong is that he failed to realize that the lower turnout in the panhandle in 2000 was really higher turnout in Democratic areas, turnout mobilized by Gore supporters. A theoretically sound view of turnout realizes that turnout is strategic, with the strategic elements controlled by mobilizers, not the mobilized.

Thus where Lott went wrong is in not thinking about turnout theoretically, not a lack of CPOs. We can see this more clearly with a more standard example, one related to studying general relationships, not explaining one specific event.

Suppose we regress Congressional vote for the incumbent on campaign spending by the incumbent. Suppose we find almost no relationship. We might conclude that money does not matter, and that everyone who thought that money did matter was wrong. This would be consistent with this regression. But no student of elections would stop there. Theory would tell us that challenger spending matters, and perhaps increased incumbent spending is related to increased challenger spending. Or perhaps incumbents in trouble spend more to offset their troubles. The electoral analyst would then incorporate these theoretical ideas (and ideas which are also consistent with knowledge of the cases) into more appropriate regressions, which would then yield more believable results. The results would be more believable because they are based on models consistent with extant theory, and because they are consistent with our knowledge of the cases. Thus no one relies (or no one should rely) on only DSOs; it is DSOs, in conjunction with theory and knowledge of the cases that allow us to obtain quantitative results that we believe. But to talk about CPOs, and to argue that one should pursue additional CPOs as opposed to additional DSOs simply does not get us very far. Understanding what DSOs might help us discriminate between alternative theories is much closer to what I think we need. But here we are back to KKV.

Note that if we wanted to understand why Incumbent X won without spending any money we might turn to a variety of explanations, some based on general laws or observed regularities, some based on idiosyncratic features. Perhaps we would talk with local experts or even journalists. But as social scientists our interest should be in generating and testing theoretical propositions, not explaining that incumbent X won without spending any money because challenger Y was caught *in flagrante delicto*.” Much of what we actually do is what KKV call “descriptive inference.” As we saw in the campaign spending example, theory can help us decide which regressions yield more valid descriptive inferences, but one should not confuse this use of theory with making observations on a causal process, let alone joining such observations to an enhanced data set.

So my take is that RSI has not really made it clear what it means to adjoin CPOs to DSOs. This undermines a variety of key points in the book, particularly as it relates to attempting to save from the wrath of KKV research designs beloved by some qualitative analysts. This should not be seen as taking away from the real merits of RSI: a much more nuanced view of the research process and a much needed examination of the cleanliness of the hands of the quantitative analyst. Both KKV and BC agree that qualitative and quantitative researchers should share the same standards; BC make a noble effort to show that the tools of the qualitative analyst meet those standards; but I find that effort not quite as persuasive as do Brady and Collier.

REFERENCES

Brady, Henry E. and David C. Collier, eds. 2004. *Rethinking Social Inquiry: Diverse Tools, Shared Standards*. Lanham, Md.: Rowman and Littlefield.

- King, Gary M., Robert O. Keohane and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton: Princeton University Press.
- 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.