Advance Access publication September 2, 2010

Political Analysis (2010) 18:506–513
doi:10.1093/pan/mpq022

Outdated Views of Qualitative Methods: Time to Move On

David Collier and Henry E. Brady
Department of Political Science, University of California, Berkeley
210 Barrows Hall #1950
Berkeley, CA 94720
e-mail: dcollier@berkeley.edu (corresponding author), hbrady@berkeley.edu

Jason Seawright
Department of Political Science, Northwestern University
Scott Hall, 601 University Place
Evanston, IL 60208
e-mail: j-seawright@northwestern.edu

Both qualitative and quantitative research routinely fall short, producing misleading causal inferences. Because these weaknesses are in part different, we are convinced that multimethod strategies are productive. Each approach can provide additional leverage that helps address shortcomings of the other. This position is quite distinct from that of Beck, who believes that the two types of analysis cannot be adjoined. We review examples of adjoining that Beck dismisses, based on what we see as his outdated view of qualitative methods. By contrast, we show that these examples demonstrate how qualitative and quantitative analysis can work together.

1 Introduction

Achieving sound causal inference is an abiding goal in political science, and we argue that adjoining qualitative and quantitative methods advances this goal. Beck's (2010) challenge to our position[1] reflects what we believe is his outdated understanding of qualitative methods. It is time to move on.

Beck questions our idea of causal process observations (CPOs),[2] suggesting that it may be an oxymoron (see also Beck 2006). He comments that "it is hard for me to know exactly what a causal-process observation is" (2010). In discussing the fundamental idea of testing theories by examining their empirical implications, Beck challenges our approach: "what

Authors' note: This response is dedicated to the statistician David A. Freedman, who thought it was late in the game for quantitative researchers still to be skeptical about qualitative methods. We received valuable suggestions and assistance from Tara Buss, F. Daniel Hidalgo, Jody LaPorte, and especially Christopher Chambers-Ju, Maria Gould, and Miranda Yaver.

[1]Our prior exchange is Beck (2006) and Brady, Collier, and Seawright (BCS, 2006).
[2]CPOs may be defined as diagnostic pieces of evidence that yield insight into causal connections and mechanisms, providing leverage for adjudicating among alternative hypotheses. CPOs are not part of a rectangular data set, and the decision to focus on particular CPOs is guided by the researcher's theoretical framework, hypotheses, and substantive knowledge—and correspondingly, by the judgment that they have strong probative value in evaluating specific explanatory claims. See Brady and Collier (2004, 252–264) (hereafter BCS 2004), BCS (2006), Freedman (2008, 300–301, 312–13), Bennett (2008), and Mahoney (2010, 123–31).

© The Author 2010. Published by Oxford University Press on behalf of the Society for Political Methodology. All rights reserved. For Permissions, please email: journals.permissions@oxfordjournals.org

https://doi.org/10.1093/pan/mpq022 Published online by Cambridge University Press

this has to do with 'causal-process observations' is beyond me.'' Later, using similar language, he states that ''good research design is fundamental. What this has to do with 'qualitative' methods is beyond me.''

Beck sees the contribution of linking CPOs and data set observations (DSOs[3]) as ''chimerical.'' He does acknowledge that ''no one could argue that knowledge of how the world works is a bad thing,'' and he recognizes that qualitative evidence can contribute to such knowledge. But in his view, this evidence cannot be joined with quantitative analysis to enhance causal inference. Beck's commitment to his particular understanding of quantitative methods is so strong that at three points, he defensively suggests that if qualitative tools can work effectively in tandem with quantitative methods, it would pose a challenge to the latter.

In contesting Beck's arguments, we first seek to level the playing field. He rejects CPOs as a basis for causal inference, whereas he is convinced that DSOs are a more credible approach. This view needs to be rebalanced. Second, we provide what we see as a telling illustration of how qualitative and quantitative analysis can indeed be adjoined: CPOs can help scholars construct stronger statistical models, which in turn improves DSO-based analysis. Finally, we examine five examples in which Beck claims CPOs do not contribute to causal inference: one strictly qualitative (Tannenwald), three involving variants of regression analysis (Lieberman, Stokes, and Brady), and one natural experiment (Snow on cholera). The examples consistently demonstrate the contribution of qualitative analysis.

2 Leveling the Playing Field

Central to Beck's claim that research based on CPOs and DSOs cannot be adjoined[4] is his view that CPOs make no contribution to causal inference, as he argues in example after example. However, he gives only fleeting attention to significant weaknesses of methods based on DSOs.[5]

The sharp criticism of regression-based causal inference by leading quantitative methodologists challenges Beck's position. These scholars argue that carrying out empirical analysis with a given statistical model does not really test the model; small changes in model specification can dramatically change the results; adding more control variables can worsen rather than improve causal inference; and among the different applications of regression, its use in causal inference is possibly the least productive (Freedman 1991, 2008; Diaconis 1998; Heckman 2000; Achen 2002; R. Berk 2004; Schrodt 2006; Seawright 2010).

Advanced refinements on regression might provide one avenue of escape, yet they may be useless or even counterproductive. Diaconis (1998) argues that such ''bells and whistles'' can ''lead to nonsense'' (797), and Freedman (1991) comments that ''technical fixes become relevant only when models are nearly right'' (305). R. Berk has grave misgivings about regression diagnostics and specification tests, and Winship and Mare (1992) express the now widely held skepticism about Heckman correction models. Parallel concerns have been raised about robust estimators, generalized least squares, and propensity scores (Freedman 1991, 305; R. Berk and Freedman 2008). Matching designs likewise do not avoid the pitfalls of statistical modeling (Sekhon 2009).

[3]That is, the quantitative analyst's observations, organized in a rectangular data set.

[4]Obviously, the question here is not whether the two forms of data can literally be merged, which would not make sense because they are simply in different formats. Rather, the question is whether they can work together in making causal inferences.

[5]Beck's conclusion does emphasize the superiority of natural experiments over regression designs. He also notes that Lott's difference-in-differences regression analysis used a poor research design, but he does not discuss this problem in light of the broader, inherent limitations of regression-based inference. For his other two examples involving regression analysis (Lieberman and Stokes), the contribution of qualitative analysis is dismissed, but the possibility of basic problems in the statistical analysis is not considered.

https://doi.org/10.1093/pan/mpq022 Published online by Cambridge University Press

Natural experiments could be an alternative avenue of escape. Although a few provide remarkably strong leverage for causal inference, as with Snow's study on cholera, many do not. In the view of Dunning (2010), this set of techniques—including regression discontinuity and instrumental variables designs—is not a panacea for scholars wishing to move beyond standard quantitative tools. First, the assumption of *as-if* random assignment essential to these designs is often not satisfied. Second, using statistical adjustments introduces again the problems of modeling. Third, the limited range of real-world situations in which these designs are feasible may drastically narrow the substantive interest and scope of their findings, thereby vitiating their contribution. Of these three concerns, the second and third also apply to many randomized experiments.

In sum, this wide spectrum of quantitative techniques based on DSOs suffers from substantial limitations. Given this conclusion, we ask: Can CPOs and DSOs possibly work together to strengthen causal inference?

3 Model Specification as Successful Adjoining

CPOs and DSOs can be productively adjoined, for example, when qualitative evidence contributes to specifying statistical models. This includes decisions about which variables to incorporate in the model and the posited relationships among them, including causal heterogeneity, functional form, and temporal sequencing.

Statistical modeling establishes the empirical relationships that hold *if* the model is correct. The empirical analysis may test some aspects of the model, such as linearity, but the validity of these limited tests depends on the correctness of the untested parts of the model.

Thus, the challenge is not just statistically *estimating* the model's parameters, but making good decisions in *constructing* the model and *justifying* its assumptions. Our two-fold argument is that CPOs can contribute to these decisions, and the basis for this contribution should be rigorous and explicit.

Let us illustrate. Qualitative evidence can bring into focus temporal discontinuities that create causal heterogeneity, which should be incorporated in the statistical model. For example, Johnston et al.'s (1992) study of the 1988 Canadian elections uses information about a dramatic confrontation between party leaders over the Canada–U.S. Free Trade Agreement, 4 weeks before the election. The simple qualitative observation of this confrontation identifies a temporal inflection point that proves crucial in the authors' quantitative analysis.

A similar example appears in Dunning's (2004) research on the impact of foreign development assistance on regime type in Africa. He argues that "while many qualitative studies have suggested that the end of the Cold War marked a watershed in the politics of aid in Africa," quantitative research has failed to account for this potential source of causal heterogeneity (410). Refining his statistical model by dividing the quantitative analysis into the Cold War and post-Cold War periods, Dunning addresses this heterogeneity and arrives at a more nuanced statistical finding.

Qualitative evidence can also contribute to decisions about the inclusion of variables—a critical choice, given that adding variables that do not belong in the underlying causal structure can harm causal inference. In Lieberman (2003), intensive case studies in Brazil and South Africa—focused on how the definition of national political community shapes state capacity to extract taxes—serve to identify a variable that is subsequently entered into his regression analysis. In Post's (2010) study of conditions under which privatization agreements succeed or collapse, fine-grained case studies in Argentina yield a parsimonious understanding of factors shaping these outcomes, which in turn helps specify the relatively simple statistical model employed in a wider quantitative test. Hidalgo et al.'s (2010) study of rural land invasions in Brazil evaluates the causal connection between

https://doi.org/10.1093/pan/mpq022 Published online by Cambridge University Press

adverse economic shocks and land seizures. Using case study evidence on causes of rural conflict, the authors add a key variable: inequality in land distribution (509–10).

Model specification is thus a domain in which CPOs and DSOs can productively work together.

4 CPOs In Fact Provide Inferential Leverage: Revisiting Beck's Examples

We now turn to the examples—involving qualitative research, regression analysis, and a natural experiment—which Beck claims demonstrate that CPOs do not contribute to causal inference. In our view, by contrast, these examples demonstrate that CPOs *do* provide inferential leverage.

4.1 *Causal Inference in Qualitative Research: Tannenwald*

Beck's discussion of Tannenwald's (1999) study of post-World War II international crises reveals his narrow perspective on causal inference and his fundamental unwillingness to consider qualitative evidence. Tannenwald evaluates a key explanation for the nonuse of nuclear weapons by the United States—a "nuclear taboo" deriving from horrified reactions to the use of nuclear weapons in World War II. Her four cases constitute a small N, with no variance on the dependent variable. Hence, treating these cases as part of a rectangular data set makes no sense, and Tannenwald turns to CPOs and process tracing, focusing on participants' accounts of their decisions.

Beck rejects Tannenwald's analysis, stating that policy makers' own accounts of their decisions may have some value, "but it is a different study from (the impossible one of) the causes of the U.S. using or not using nuclear weapons after World War II."

Two things strike us here. First, it is wrong to claim that it is "impossible" to study the nonuse of nuclear weapons (i.e., a nonevent) merely because they were not employed during a given period. With such claims, methodology simply buries its head in the sand.

Second, Beck dismisses the accounts of policy makers: "sometimes they tell stories we like, and we are happy, and sometimes not." Yet, the same is true of regression coefficients: sometimes they make us happy, sometimes not, and they are not necessarily a reliable basis for causal inference. Beck's criticism of Tannenwald can equally well be applied to quantitative methods.

4.2 *Adjoining CPOs with Regression Analysis: Lieberman, Stokes, and Brady*

Beck objects to our analysis of Lieberman (2003), arguing that by including him in the discussion, we stretch the idea of adjoining. However, CPOs and DSOs do work together here. As noted above, Lieberman's case studies help him specify his regression model, which we view as a fundamental contribution of qualitative evidence to quantitative research.

Stokes (2001) focuses on newly elected presidents in Latin America who campaign on antineoliberal promises and then, postelection, shift to neoliberal policies. She combines cross-national quantitative analysis with qualitative case studies to evaluate the causes of the shift. Here, Beck again refuses to give credence to causal inference based on qualitative analysis. Yet, one could just as well be skeptical of Stokes's use, in her quantitative analysis, of the Heckman selection model (e.g., 95–99) about which some scholars have serious misgivings. Why does Beck insist that it is specifically the qualitative part of Stokes's analysis—rather than the quantitative/Heckman part—that yields unconvincing inferences?

Finally, Beck's treatment of Brady (2004) puzzles us. Brady seeks to demonstrate the inferential power of CPOs, whereas Beck maintains that "there was nothing in Brady's discussion that was not in the standard quantitative analyst's toolkit." Beck does not accept the fact that Brady analyzes CPOs.

https://doi.org/10.1093/pan/mpq022 Published online by Cambridge University Press

The substantive issue is the following (see BCS 2004, 267–71). For the 2000 presidential election in Florida, Lott hypothesizes that the early media call—which incorrectly declared a Gore victory—strongly affected the Bush vote in the Panhandle. The Panhandle (unlike the rest of Florida) is on Central Time, so that many voters might have heard the call before polls closed. Using a difference-in-differences regression analysis, Lott concludes that Bush lost at least 10,000 votes in the Panhandle.

Brady's working hypothesis, by contrast, is that the early media call had little or no effect. He constructs a sequence of tests based on CPOs to estimate how many voters could have been influenced by the early call. For claims regarding a large causal effect to be credible, these estimates must have high values. Instead, Brady finds that these values are small. Brady's analysis is not based on DSOs. It involves a series of conjectures and pieces of evidence about this particular case rather than a systematic scoring of this case in relation, for example, to other areas in Florida. Based on these tests, he concludes that the early call could *not* have significantly suppressed the Bush vote.

This example is valuable for three reasons. It is a reminder that in addition to being routinely and appropriately seen as an aspect of qualitative methods, CPOs can also involve numerical data (see also Freedman 2008, 300–301). The difference is that the numbers are not analyzed as part of a rectangular data set. This distinction is crucial to our disagreement with Beck's statement that Brady conducted only standard quantitative analysis.

Brady's study also shows that the diagnostic power of CPOs can be enhanced by background information derived from quantitative research. This is a standard and appropriate approach to clarifying the context in which a CPO is situated, and it does not make the study quantitative. Thus, in our view—and contra Beck—Brady's analysis is an excellent example of causal inference that distinctively relies on CPOs.

Finally, Beck argues that it makes no sense that CPOs should be adjoined with DSOs to resolve the dispute over the vote loss. Clearly, this is a different form of adjoining than in the other examples, where the two types of observations work together in a study conducted by one researcher, with a single research design. In the Florida analysis, by contrast, CPOs are used to critique an inference, based on DSOs, that is advanced by another scholar. This is a distinct mode of adjoining, but an instance of it nonetheless.

4.3 *Adjoining CPOs with a Natural Experiment: Snow on Cholera*

Beck poses the question "Who Gets to Claim John Snow?"—referring to the brilliant epidemiologist who demonstrated that cholera is a waterborne disease. Beck invokes Snow to illustrate his argument that CPOs and quantitative analysis cannot work together in achieving causal inference. We think that Beck is wrong and instead share Freedman's (2008) conclusion that Snow exemplifies the contribution of CPOs.

We agree with Beck that Snow's natural experiment—which Beck calls a quasiexperiment[6]—was remarkable. Dunning (2010) reaches the same conclusion, while also demonstrating that many such designs are far less successful. Given that Snow's study is indeed exceptional, this should be a particularly likely example for supporting Beck's main argument. If the design is so powerful, why should qualitative evidence be relevant?

[6]Beck consistently uses the expression *quasiexperiment* in referring to Snow. We recommend against his usage, strongly preferring *natural experiment*. This latter term is usefully understood as a family of techniques that employ the assumption of *as-if* random assignment and thereby includes regression discontinuity and instrumental variables designs. The term quasiexperiment was popularized by Campbell and Stanley (1963) and Campbell and Ross (1968) to mean interrupted time-series designs. Here, the key assumption of *as-if* random assignment is not appropriate. Later in his career, Campbell expressed regret that he had promoted the idea of quasiexperiments because the term created a false impression of experimental rigor (BCS 2004, 230–31).

https://doi.org/10.1093/pan/mpq022 Published online by Cambridge University Press

We think that Beck draws the wrong lesson from Snow. Although Beck acknowledges that some qualitative information was helpful in Snow's research, he argues that it was not central to the investigation. Beck believes that Snow's work ''shows 'only' excellent quantitative analysis rather than any challenges to such analysis,'' which in our view reveals Beck's defensive posture. In response, we ask: If qualitative evidence proves important, would it truly challenge quantitative methods? We cannot understand why. Qualitative methods would open an avenue for strengthening—not challenging—quantitative analysis.

Freedman (1991, 2008), who introduced Snow into debates on social science methodology, offers an interpretation radically different from Beck. Specifically using the language of CPOs, Freedman (2008, 301, 313) argues that Snow's findings (along with several other breakthrough studies in epidemiology, which generally incorporated DSOs into their research designs) depended on joining the leverage provided by quantitative and qualitative methods. Freedman seeks to demonstrate ''the role played by CPOs, and by qualitative reasoning more generally'' in innovative research. Snow's analysis ''illustrates the power of case studies'' (305). Indeed, this use of qualitative evidence deserves recognition as a ''type of scientific inquiry'' in its own right (301):

> progress depends on refuting conventional ideas if they are wrong, on developing new ideas that are better, and on testing the new ideas as well as the old ones. The examples show that qualitative methods can play a key role in all three tasks. (Freedman 2008, 312)

Freedman thus asserts that CPOs perform a full range of inferential tasks. It is unfortunate that Beck would see this teamwork between qualitative and quantitative tools as a challenge to quantitative methods. We view it instead as precisely that: teamwork.

The details of Snow's study are crucial here. Early on, he abandoned the established explanation that cholera spread through '''miasmas,' that is, noxious odors—especially odors generated by decaying organic material'' (Freedman 2008, 304). Instead, he moved toward the idea that cholera spread through human contact and was waterborne. Snow observed that cholera was passed along lines of commerce. He saw that the first case in London involved the death of a sailor just arrived from Hamburg, where there was a cholera epidemic. The second was an individual who subsequently slept in the same boarding house room as this sailor.

Snow was ''intimately familiar with the Broad Street area [the center of the epidemic], because of his medical practice'' (305), which contributed to astute inferences about the incidence of cholera in particular neighborhoods, workplaces, and households. Snow also observed that two water companies served this area—one apparently providing contaminated water and the other not; and he recognized that households receiving water from the two companies were intermingled on the same streets. These insights provided the basis for assuming that allocation of contaminated water occurred *as-if* at random, thus justifying the natural experiment. It is crucial for our argument that Snow arrived at these insights *prior to* the confirmation provided by this experiment (Freedman 2008, 306).

There is no disagreement that Snow's natural experiment yielded an unusually decisive confirmation of his main hypothesis (Freedman 2008, 306). Yet, Beck fails to acknowledge that the construction of this remarkable study was heavily dependent on CPOs.

Snow's research shows how constructing a natural experiment is in many ways like constructing a statistical model. In both cases, qualitative evidence and CPOs can play critical roles in strengthening DSOs. We therefore definitely ''claim John Snow'' to support our position on adjoining.

https://doi.org/10.1093/pan/mpq022 Published online by Cambridge University Press

5 Conclusions

In his conclusion, Beck seeks to establish common ground with us, which we applaud. Yet we question his version of common ground. He sharply dismisses CPOs and the possibility that CPOs and DSOs can work together, reflecting once again what we see as his outdated understanding of qualitative methods.

In Beck's view, our disagreement is just a question of "nomenclature"; he suggests that we simply drop the terms qualitative and quantitative. We too sometimes find these terms confining, but we insist that CPOs and DSOs provide distinctive inferential leverage. This is not an issue of names and labels.

Beck shares our admiration for Snow's natural experiment, yet his comments on Snow distract from the main point under discussion. Beck conspicuously neglects to mention our own argument—strongly supported by Freedman—that qualitative evidence was essential to Snow's construction of the research design. Furthermore, Dunning's comprehensive evaluation of natural experiments emphasizes more broadly the centrality of qualitative evidence in formulating such designs. This is a story about adjoining.

Thus, to reiterate, we appreciate Beck's effort to establish common ground. Yet, the idea of common ground is plausible only with teamwork between qualitative and quantitative methods, based on recognition that they can jointly enhance causal inference.

References

Achen, Christopher H. 2002. Toward a new political methodology: Microfoundations and ART. *Annual Review of Political Science* 5:423–50.

Beck, Nathaniel. 2006. Is causal-process observation an oxymoron? *Political Analysis* 14:347–52.

———. 2010. Causal process "observation": Oxymoron or (fine) old wine? *Political Analysis* 18(4):499–505.

Bennett, Andrew. 2008. Process tracing: A Bayesian perspective. In *Oxford handbook of political methodology*, ed. J. Box-Steffensmeier, H. E. Brady, and D. Collier, 702–21. Oxford: Oxford University Press.

Berk, Richard A. 2004. *Regression analysis: A constructive critique*. Thousand Oaks, CA: Sage Publications.

Berk, Richard A., and David A. Freedman. 2008. On weighting regressions by propensity scores. *Evaluation Review* 32:392–409.

Brady, Henry E. 2004. Data-set observations versus causal-process observations: The 2000 U.S. presidential election. In *Rethinking social inquiry: Diverse tools, shared standards*, ed. H. E. Brady and D. Collier, 267–71. Lanham, MD: Rowman & Littlefield.

Brady, Henry E., and David Collier. 2004. *Rethinking social inquiry: Diverse tools, shared standards*. Lanham, MD: Rowman & Littlefield. Given Jason Seawright's coauthorship of many chapters, the in-text reference is BCS 2004.

Brady, Henry E., David Collier, and Jason Seawright. 2006. Toward a pluralistic vision of methodology. *Political Analysis* 14:353–68.

Campbell, Donald T., and H. Laurence Ross. 1968. Connecticut crackdown on speeding: Time-series data in quasi-experimental analysis. *Law and Society Review* 3(1):33–53.

Campbell, Donald T., and Julian C. Stanley. 1963. *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.

Diaconis, Persi. 1998. A place for philosophy? The rise of modeling in statistical science. *Quarterly of Applied Mathematics* 56:797–805.

Dunning, Thad. 2004. Conditioning the effects of aid: Cold War politics, donor credibility, and democracy in Africa. *International Organization* 58:409–23.

———. 2010. Design-based inference: Beyond the pitfalls of regression analysis? In *Rethinking social inquiry: Diverse tools, shared standards*. 2nd ed., ed. H. E. Brady and D. Collier, 273–311. Lanham, MD: Rowman & Littlefield.

Freedman, David A. 1991. Statistical models and shoe leather. *Sociological Methodology* 21:291–313.

———. 2008. On types of scientific inquiry: The role of qualitative reasoning. In *Oxford handbook of political methodology*, ed. J. Box-Steffensmeier, H. E. Brady, and D. Collier, 300–18. Oxford: Oxford University Press.

Heckman, James J. 2000. Causal parameters and policy analysis in economics: A twentieth century retrospective. *Quarterly Journal of Economics* 115:45–97.

https://doi.org/10.1093/pan/mpq022 Published online by Cambridge University Press

Hidalgo, F. Daniel, Suresh Naidu, Simeon Nichter, and Neal Richardson. 2010. Economic determinants of land invasions. *Review of Economics and Statistics* 92(3):505–23.

Johnston, Richard, André Blais, Henry E. Brady, and Jean Crête. 1992. *Letting the people decide: Dynamics of a Canadian election*. Stanford, CA: Stanford University Press.

Lieberman, Evan S. 2003. *Race and regionalism in the politics of taxation in Brazil and South Africa*. New York: Cambridge University Press.

Mahoney, James. 2010. After KKV: The new methodology of qualitative research. *World Politics* 62(1):120–47.

Post, Alison. 2010. *Liquid assets and fluid contracts: Explaining the uneven effects of water and sanitation privatization*. Berkeley, CA: Department of Political Science, University of California.

Schrodt, Philip. 2006. Beyond the linear frequentist orthodoxy. *Political Analysis* 14:335–39.

Seawright, Jason. 2010. Regression-based inference: A case-study in failed causal assessment. In *Rethinking social inquiry*. 2nd ed., ed. H. E. Brady and D. Collier, 247–71. Lanham, MD: Rowman & Littlefield.

Sekhon, Jasjeet. 2009. Opiates for the matches: Matching methods for causal inference. *Annual Review of Political Science* 12:487–508.

Stokes, Susan. 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 for nuclear non-use. *International Organization* 53:433–68.

Winship, Christopher, and Robert D. Mare. 1992. Models for sample selection bias. *Annual Review of Sociology* 18:327–50.

https://doi.org/10.1093/pan/mpq022 Published online by Cambridge University Press