Comments on "Covariance Adjustments for the Analysis of Randomized Field Experiments"

Winston Lin[1]

Richard Berk, Emil Pitkin, Lawrence Brown, Andreas Buja, Edward George, and Linda Zhao (2014) have written a valuable *Evaluation Review* paper on regression adjustment in randomized experiments. I've long been a fan of Berk's critical writings on regression and meta-analysis (Berk 2004, 2007; Berk and Freedman 2003), and I recently recommended Berk, Brown, et al.'s (2014) helpful *Sociological Methods and Research* paper to a colleague teaching a course on regression. Also, I am grateful to Berk for sending me kind and constructively critical comments on Lin (2013) after it went to press. In my reply to his e-mail, I shared an informal essay (Lin 2012a, 2012b) discussing regression adjustment in practice. We are in agreement on many points.

My comments below are only to clarify a few of Berk, Pitkin, et al.'s (2014) statements about Lin (2013):

1. "He replicates Freedman's overall results and then turns to a conceptual framework that differs substantially from Freedman's." (Berk, Pitkin, et al. 2014, p. 172)

2. "For Lin, the population from which the subjects are drawn is real and finite. The researcher is assumed to know the population mean for the covariate, which can be used as the value of θ. In most social science applications, that mean will not be known." (Berk, Pitkin, et al. 2014, p. 179)

3. "Lin implicitly loosens the ties to the Neyman approach by making use of a real, finite population from which the data can be treated as a random sample. His conclusions are less pessimistic than Freedman's. However, his proposed estimator will usually not be operational in practice …" (Berk, Pitkin, et al. 2014, p. 187)

In fact, I stuck rigidly to Neyman's and Freedman's framework: "The *n* subjects are the population of interest; they are not assumed to be randomly drawn from a superpopulation" (Lin 2013, p. 297). Since the "population" is just the *n* subjects in the experiment, the population mean for the covariate is known, and the OLS-with-interactions estimator that I studied is identical to Berk, Pitkin, et al.'s (2014, p. 182, Equation 11) estimator. Their paper helpfully provides a different standard error estimator (Berk, Pitkin, et al. 2014, p. 183, Equation 12) because they are generalizing to an infinite population. (I am agnostic about finite- vs. infinite-population inference. In my paper, I recommended Reichardt and Gollob's [1999, pp. 125–127] discussion and wrote, "My purpose is not to advocate finite-population inference, but to show just how little needs to be changed to address Freedman's major concerns.")

---

[1] Winston Lin (Ph.D., Statistics, University of California, Berkeley, 2013) will be a postdoctoral research scholar in the Department of Political Science at Columbia University, beginning in September 2014. E-mail: Linston@gmail.com.

Perhaps the confusion stems from the "imaginary infinite sequence of finite populations" (Lin 2013, p. 301). This is merely the same setup that Freedman assumed for his asymptotic results (and my regularity conditions are merely Freedman's [2008b] regularity conditions, generalized to allow multiple covariates). He writes (Freedman 2008a, p. 184), "In principle, our inference problem is embedded in an infinite sequence of such problems, with the number of subjects n increasing to infinity." However, the goal here is still to make inferences about the average treatment effect on the actual subjects in the experiment. As Lehmann (1999, p. 255) writes: "We must go back to the purpose of embedding a given situation in a fictitious sequence: to obtain a simple and accurate approximation. The embedding sequence is thus an artifice and has only this purpose ..."

These are but small points in the larger scheme of things, and I look forward to much useful research from Berk, Pitkin, Brown, Buja, George, and Zhao.

References

Berk, R. A. 2004. *Regression Analysis: A Constructive Critique*. Thousand Oaks, CA: Sage.

Berk, R. A. 2007. "Statistical Inference and Meta-Analysis" (with discussion). *Journal of Experimental Criminology* 3:247–297.

Berk, R. A., L. Brown, A. Buja, E. George, E. Pitkin, K. Zhang, and L. Zhao. 2014. "Misspecified Mean Function Regression: Making Good Use of Regression Models That Are Wrong." *Sociological Methods and Research* 43:422–451.

Berk, R. A., and D. A. Freedman. 2003. "Statistical Assumptions as Empirical Commitments." In T. G. Blomberg and S. Cohen (Eds.), *Punishment and Social Control*, 2nd ed., 235–254. New York: de Gruyter.

Berk, R. A., E. Pitkin, L. Brown, A. Buja, E. George, and L. Zhao. 2014. "Covariance Adjustments for the Analysis of Randomized Field Experiments." *Evaluation Review* 37:170–196.

Freedman, D. A. 2008a. "On Regression Adjustments to Experimental Data." *Advances in Applied Mathematics* 40:180–193.

Freedman, D. A. 2008b. "On Regression Adjustments in Experiments With Several Treatments." *Annals of Applied Statistics* 2:176–196.

Lehmann, E. L. 1999. *Elements of Large-Sample Theory*. New York: Springer.

Lin, W. 2012a. "Regression Adjustment in Randomized Experiments: Is the Cure Really Worse Than the Disease? (Part I)" [Blog post]. Retrieved from
http://blogs.worldbank.org/impactevaluations/node/847

Lin, W. 2012b. "Regression Adjustment in Randomized Experiments: Is the Cure Really Worse Than the Disease? (Part II)" [Blog post]. Retrieved from http://blogs.worldbank.org/impactevaluations/node/849

Lin, W. 2013. "Agnostic Notes on Regression Adjustments to Experimental Data: Reexamining Freedman's Critique." *Annals of Applied Statistics* 7:295–318.

Reichardt, C. S., and H. F. Gollob. 1999. "Justifying the Use and Increasing the Power of a *t* Test for a Randomized Experiment With a Convenience Sample." *Psychological Methods* 4:117–128.