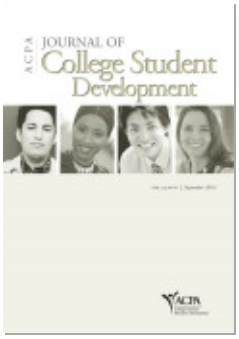
PROJECT MUSE®

Design and Analysis in College Impact Research: Which Counts More?



Ernest T. Pascarella, Mark H. Salisbury, Charles Blaich

Journal of College Student Development, Volume 54, Number 3, May-June 2013, pp. 329-335 (Article)

Published by The Johns Hopkins University Press
DOI: 10.1353/csd.2013.0032

For additional information about this article
http://muse.jhu.edu/journals/csd/summary/v054/54.3.pascarella.html

Access provided by Penn State Univ Libraries (19 Sep 2015 12:19 GMT)

Design and Analysis in College Impact Research: Which Counts More?

Ernest T. Pascarella Mark H. Salisbury Charles Blaich

Over the last several decades student affairs and assessment scholars who study college impact have utilized a number of different research design and statistical procedures in an attempt to control for the characteristics and propensities that lead students to self-select themselves into a particular intervention or experience. This is particularly important because such characteristics and propensities may seriously confound any estimate of the effect of the intervention or experience itself. By far the most common method used in the college impact literature to date has been covariate adjustment, based on different multiple regression approaches (Pascarella & Terenzini, 2005). This approach relies on statistical control to remove or partial out the confounding effects of student self-selection. Recently, however, there has been considerable criticism of covariate adjustment based on the argument that its estimate of the effect of an intervention or experience can be biased. Rather than relying on regression-based covariate adjustment techniques, a number of scholars have suggested the use of propensity score matching as a more effective analytical approach for controlling the effects of demographic, attitudinal, or other factors that might increase or decrease students' likelihood of self-selecting into a given treatment of interest and, thereby, isolating the effect of the treatment itself (e.g., Reynolds & DesJardins, 2009; Schneider, Carnoy, Kilpatrick, Schmidt, & Shavelson, 2007).

In this study we employ both covariate adjustment and propensity score matching to estimate the causal influence of an example intervention—the first year of attendance at a liberal arts college (as opposed to another type of 4-year institution). Specifically we estimated the effect of liberal arts college attendance on three cognitive outcomes. We examined the estimates yielded by these two analytical approaches under different research design assumptions—with and without a precollege measure of each outcome. Our purposes were to determine the comparability of causal estimates using covariate adjustment and propensity score matching, and to examine how these estimates might be affected when different research designs are employed to study college impact. The focus of the study was not specifically on understanding the effects of liberal arts colleges, rather, estimating of the effects of liberal arts colleges versus other 4-year institutions is used only as an example. The approaches we explored could have relevance to estimating of the effects of a broad range of between-college and within-college interventions or experiences.

Ernest T. Pascarella is Professor and Mary Louise Petersen Chair in Higher Education at The University of Iowa. Mark H. Salisbury is Assistant Dean and Director of Institutional Research and Assessment at Augustana College. Charles Blaich is Director of the Center of Inquiry in the Liberal Arts at Wabash College. *Research supported by a generous grant from the Center of Inquiry in the Liberal Arts at Wabash College to the Center for Research on Undergraduate Education (CRUE) at The University of Iowa. A substantially more detailed version of this paper is available from http://www.education.uiowa.edu/crue*

METHODS

Sample and Data Collection

We analyzed data from the first year of the Wabash National Study of Liberal Arts Education (WNS), which is a longitudinal pretest–posttest investigation of the effects of liberal arts experiences on a range of college cognitive and noncognitive outcomes thought to be associated with undergraduate liberal arts education. The colleges and universities participating in WNS represent a diverse selection of institutions, varying in institutional characteristics such as type and control, selectivity, enrollment, and location within the United States. For our data analysis sample we chose the 2006 WNS iteration, which collected extensive precollege data on students in early Fall 2006 and again in Spring 2007. Our analyses were based on first-year, full-time undergraduates attending 17 different 4-year institutions (11 liberal arts colleges, 3 research universities, and 3 regional institutions). We estimated the effects of the first year of liberal arts college attendance on three cognitive/learning orientation outcome measures: critical thinking skills, need for cognition (a measure of continuing motivation for learning), and positive attitude toward literacy activities. Because of matrix sampling in part of the WNS design, complete precollege and end-of-first-year data were available for 1,377 students on one dependent measure (critical thinking skills) and 2,872 students on the other two dependent measures (need for cognition and positive attitude toward literacy). Although there are clearly limitations with respect to the external validity or generalizability of results obtained with the 17-institution WNS sample, our concern was with estimating the internal validity of the effects of liberal arts colleges. Moreover, as our results are intended for didactic rather than inferential purposes, concerns with generalizing the results of the analyses are largely irrelevant to our purpose.

Independent and Selection Effect Variables

The independent (or treatment) variable in the study was a simple dummy variable (i.e., 1/0) where 1 = attendance at a liberal arts college, and 0 = attendance at another type of 4-year institution (i.e., a research university or regional institution). Central to the study, however, was the fact that significantly different kinds of students were likely to self-select themselves into liberal arts colleges and other types of 4-year institutions (Pascarella, Wolniak, Seifert, Cruce, & Blaich, 2005; Pascarella & Terenzini, 2005). At the inception of the WNS, surveyors were cognizant of this selection effect and made a purposeful effort to gather a wide range of precollege characteristics on each student participant in the study. Our first step in taking the selection effect into account was to examine this precollege data to determine which of these characteristics was a significant predictor of attendance at a liberal arts college versus another type of 4-year school. This produced a list of 12 variables, most of which seemed quite reasonable given prior evidence on who goes where to college (Pascarella & Terenzini, 2005). In addition to race and sex, which were included because of some sample bias, these variables then became the basis for constructing our matching and covariate adjustment models. The selection effect variables were as follows: race, sex, parental education, ACT composite score, receipt of a federal grant to attend college (a proxy for SES), receipt of an institutional grant to attend college, precollege intent to major in a preprofessional field, precollege political views, precollege scores on a purpose in life scale, precollege critical thinking test scores, precollege need for cognition scores, and precollege attitude toward literacy scores. (Operational definitions of all selection and dependent variables are available from the

expanded version of this article on the CRUE web site, or from the contact author.)

Data Analyses

We employed two analytical approaches to estimate the unique effects of the first year of attendance at a liberal arts college versus another type of 4-year institution on our three cognitive/learning orientation outcomes. In the first approach we used a propensity score matching method to directly account for the likelihood of sample selection bias. We utilized the PSMATCH2 module (Leuven & Sianesi, 2003), which is available as part of the STATA (version 11) statistical software package. The PSMATCH2 module allows users to select from a range of propensity score matching procedures to examine differences between the treated and untreated groups visually and statistically, generate an appropriate propensity score, choose from a variety of algorithms to match comparable cases using the propensity score, and estimate an average treatment effect. We employed the radius approach to matching, which allowed us to match treated cases (students attending liberal arts colleges) with all untreated cases (students attending another type of 4-year institution) within a caliper of 0.1. Employing the single PSMATCH2 command, we generated a propensity score estimating the likelihood of attending a liberal arts college using all the covariates or selection effect variables, and then used the radius matching method to estimate the average effect of attending a liberal arts college. The second analytic approach was the somewhat more familiar covariate adjustment, which is based on ordinary least squares regression. In this approach we regressed each of the three dependent measures on the dummy variable representing attendance at a liberal arts college versus another type of 4-year institution plus the same selection effect variables (covariates) used in developing the propensity scores.

In both propensity score matching and covariate adjustment approaches we estimated two different models and compared them to each other and to an unadjusted model (simple, unadjusted mean differences between students at liberal arts colleges and their counterparts at other 4-year institutions). In the first models we estimated the liberal arts college effect with propensity score matching and covariate adjustment based on all selection effect variables/covariates except a parallel pretest for each of the three end-of-first-year outcome variables. In the second models we simply added the pretest to the first models.

RESULTS

As expected, our selection effect variables as a group were strong predictors of liberal arts college attendance. Only two of the precollege variables significantly predicting attendance at a liberal arts college—parents' education and receipt of a federal grant for college attendance—were not significant unique predictors in the combined model. In short, there appeared to be a pronounced selection effect involved in attending a liberal arts college versus other types of 4-year institutions.

Table 1 summarizes the estimated effects of liberal arts college attendance on the three end-of-first-year outcomes: critical thinking skills, need for cognition, and positive attitude toward literacy. As row 1 of Table 1 shows, the unadjusted differences between liberal arts college students and their counterparts at other types of 4-year colleges were all statistically significant. Rows 2 and 3 in Table 1 summarize the estimates of liberal arts college effects taking into account all selection effect variables except the parallel precollege (or pretest) score on the outcome. In both propensity score matching and covariate adjustment models the effects of liberal arts colleges (both positive and negative) are reduced substantially for

TABLE 1.

Estimated Effects of Liberal Arts Colleges on First-Year Critical Thinking, Need for Cognition, and Positive Attitude Toward Literacy

	Critical Thinking Skills		Need for Cognition		Positive Attitude Toward Literacy	
Model (Row)	**Coefficient (*SE*)**	***t* Value**	**Coefficient (*SE*)**	***t* Value**	**Coefficient (*SE*)**	***t* Value**
1 . Unadjusted Estimate	–.235 (.054)	–4.34**	.124 (.037)	3.31**	.183 (.037)	4.34**
2 . Propensity Score Matching Estimate Without Pretest[a]	–.098 (.066)	–1.49	.119 (.045)	2.63**	.125 (.046)	2.72**
3 . Covariate-Adjusted Estimate Without Pretest[a]	–.120 (.043)	–2.79**	.134 (.034)	3.91**	.123 (.036)	3.45**
4 . Propensity Score Matching Estimate With Pretest[b]	–.039 (.066)	–0.60	.071 (.046)	1.55	.075 (.046)	1.62
5 . Covariate-Adjusted Estimate With Pretest[b]	–.045 (.037)	–1.23	.076 (.027)	2.81**	.067 (.028)	2.42*

[a] Propensity score matching and covariate adjustment based on all the following variables, except a parallel pretest, for each outcome: race, sex, parental education, ACT composite score, federal grant, institutional grant, precollege intent to major in a preprofessional field, precollege political views, precollege purpose in life, plus precollege need for cognition and positive attitude toward literacy in the prediction of end-of-first-year critical thinking, precollege positive attitude toward literacy in the prediction of end-of-first-year need for cognition, and precollege need for cognition in predicting end-of-first-year positive attitude toward literacy.

[b] Propensity score matching and covariate adjustment based on all relevant variables in note a above, plus a parallel pretest for each end-of-first-year outcome.

*$p < .05$. **$p < .01$.

critical thinking and positive attitude toward literacy. However, for end-of-first-year need for cognition the estimated effects of liberal arts colleges are essentially unchanged.

Rows 4 and 5 in Table 1 summarize the estimated effects of liberal arts college attendance on all three end-of-first-year outcomes when the pretest is also taken into account. Irrespective of the use of propensity score matching or covariate adjustment, the inclusion of the pretest substantially reduced the estimates of liberal arts college effects from the corresponding effects estimated with all selection effects variables, except the pretest (i.e., rows 2 and 3 in Table 1). Moreover, as presented in rows 4 and 5 in Table 1, while the effects from propensity score matching are slightly smaller than those from covariate adjustment for critical thinking and need for cognition, and slightly larger for positive attitude toward literacy, the two analytical approaches produced estimated effects of liberal arts college attendance that were essentially of the same magnitude. Indeed, across all three dependent measures in Table 1, the most stringent model (i.e., including a pretest) reduced bias in estimated liberal arts college effects an average of 61.7% with propensity score matching, i.e., (row 1 coefficient – row 4 coefficient)/row 1 coefficient; and an average of 61.0% with covariate adjustment, i.e., (row 1 coefficient – row 5 coefficient)/row 1 coefficient. The difference in statistical significance between propensity score matching and covariate adjustment approaches can be attributed to the increased statistical power with covariate adjustment due to reduction in the error term which leads to smaller standard errors.

Of particular importance is the role of the pretest in reducing bias in both propensity score matching and covariate adjustment approaches. Across all three outcomes in Table 1 the addition of the pretest to models that included all selection effect variables except the pretest reduced the propensity score matching estimate of liberal arts college effects an average of 46.8%, i.e., (row 2 coefficient – row 4 coefficient)/row 2 coefficient; and the corresponding covariate adjustment effects an average of 50.4%, i.e., (row 3 coefficient – row 5 coefficient)/row 3 coefficient. In short, both analytic approaches substantially overestimated the effects of liberal arts colleges when the pretest was not taken into account.

CONCLUSIONS AND DISCUSSION

The first purpose of the study was to observe the similarity in estimated effect sizes yielded by propensity score matching and covariate adjustment. Overall, when the most conservative model was estimated (taking into account a precollege measure of the outcome) the two analytical approaches produced liberal arts college effects that were remarkably similar, if not identical, in magnitude—and reduced bias almost exactly the same amount. Thus, there was little evidence in our study to suggest that propensity score matching provided an appreciably more accurate estimate of liberal arts college effects than did more traditional regression-based covariate adjustment. This is consistent with other recent evidence concerning the comparability of propensity score matching and covariate adjustment estimates (e.g., Shadish, Clark, & Steiner, 2008; Shah, Laupacis, Hux, & Austin, 2005). The choice of analytic method may not matter appreciably; however, as propensity score matching routines in statistical packages become more readily available and user friendly, it may be a prudent course of action to estimate intervention effects with both covariate adjustment and propensity score matching.

The second and perhaps more important purpose of the study was to determine how differences in research design affect the

estimates of liberal arts college effects on the three outcomes. We found that both statistical methods seriously overestimated the effects of attending a liberal arts college when our analyses did not include a parallel precollege measure (or pretest) of the outcome among the covariates / selection effect variables. Adding a pretest measure to the models reduced bias in the estimated effects of liberal arts college attendance 46.8% using propensity score matching and 50.4% using covariate adjustment—above and beyond the corresponding estimates yielded when all covariate / selection effect variables, except the pretest, were included in the models. Thus, without the inclusion of a precollege measure of each outcome, both propensity score matching and covariate adjustment would appear to have substantially overestimated the unique effects of first-year attendance at a liberal arts college. Similar findings with the WNS resulted from estimates of the effect of volunteer participation on psychological well-being (Padgett, Salisbury, An, & Pascarella, 2011) and for a broad range of other first-year outcomes (Seifert, Pascarella, Erkel, & Goodman, 2011). Such evidence underscores the salient, and arguably indispensible, role that a longitudinal, pretest–posttest design plays in estimating the developmental impacts of college. Yet, as indicated by Seifert et al. (2011) only an estimated 25% of college impact studies in four major higher education journals during a recent 5-year period employed a pretest–posttest longitudinal design.

Research design would seem to have been a more important consideration in obtaining accurate estimates of liberal arts college effects than the choice of a statistical analysis method. Design would appear to trump analysis (Rubin, 2008) and, when it is feasible, a precollege measure of any outcome may be indispensible in obtaining the most internally valid estimates of college effects. Put another way, irrespective of the analytical approach selected, estimates of causal effects in college impact research that are based on cross-sectional designs, or even longitudinal designs without a precollege measure of the outcome, may run the risk of misinterpretation—and in some cases rather serious misinterpretation.

Our conclusions may have implications for student affairs professionals charged with assessing the value-added impact of educational interventions or other college experiences. Obviously it is desirable to get timely assessment information to policy makers, and cross-sectional studies without a precollege measure of an outcome can provide a reasonable first step in producing such information. However, if the ultimate goal of student affairs assessment is to present policy makers with an accurate estimate of the causal impact of an intervention or experience, there may simply be no acceptable shortcut or alternative to pretest–posttest designs that follow participants and nonparticipants over time.

Of course just how important it is to take a pretest into account may well depend on how strongly it is associated with not only the posttest or dependent variable but also with selection into the intervention under consideration. If a pretest exists, the extent to which its inclusion reduces bias in estimating an intervention's effect is essentially an empirical question. With no pretest, however, getting the answer to this question would appear to be largely speculation; and estimates of an intervention's effect based on analyses without a pretest always have the potential for bias in unknown ways. Thus, when at all possible, the most prudent approach would be to consider a pretest–posttest longitudinal design as the research design of choice in estimating the causal influence of any college experience.

REFERENCES

Leuven, E., and Sianesi, B. (2003). *PSMATCH2: Stata module to perform full mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing.* Retrieved from http://ideas.repec.org/c/boc/bocode/s432001.html

Padgett, R., Salisbury, M., An, B., & Pascarella, E. (2011). Required, practical or unnecessary? An examination of propensity score matching using longitudinal data. *New Directions for Institutional Research,* (pp. 29-42). San Francisco, CA: Jossey-Bass.

Pascarella, E., & Terenzini, P. (2005). *How college affects students (Vol. 2): A third decade of research*. San Francisco, CA: Jossey-Bass.

Pascarella, E., Wolniak, G., Seifert, T., Cruce, T., & Blaich, C. (2005). *Liberal arts colleges and liberal arts education: New Evidence on impacts.* San Francisco, CA: Jossey-Bass.

Reynolds, C. L., & DesJardins, S. L. (2009). The use of matching methods in higher education research: Answering whether attendance at a 2-year institution results in differences in educational attainment. In J. C. Smart (Ed.), *Higher education: Handbook of theory and research*, (pp. 47-104). New York, NY: Springer Science + Business Media.

Rubin, D. (2008). For objective causal inference, design trumps analysis. *Annals of Applied Statistics, 2,* 808-840.

Seifert, T., Pascarella, E., Erkel, S., & Goodman, K. (2011). The importance of longitudinal pretest–posttest designs in estimating college impact. *New Directions for Institutional Research,* (pp. 5-16). San Francisco, CA: Jossey-Bass.

Schneider, B., Carnoy, M., Kilpatrick, J., Schmidt, W. H., & Shavelson, R. J. (2007). *Estimating causal effects using experimental and observational designs* (report from the Governing Board of the American Educational Research Foundation Grants Program). Washington, DC: American Educational Research Association.

Shadish, W. R., Clark, M. H., & Steiner, P. M. (2008). Can nonrandomized experiments yield accurate answers? A randomized experiment comparing random and nonrandom assignments. *Journal of the American Statistical Association, 103,* 1334-1356.

Shah, B., Laupacis, A., Hux, J., & Austin, C. (2005). Propensity score methods gave similar results to traditional regression modeling in observational studies: A systematic review. *Journal of Clinical Epidemiology, 56,* 550-559.