Accounting, Organizations and Society 39 (2014) 545–549

Contents lists available at ScienceDirect

Accounting, Organizations and Society

journal homepage: www.elsevier.com/locate/aos

Strengthening causal inferences in positivist field studies

Christopher D. Ittner *

The Wharton School, University of Pennsylvania, 1326 Steinberg Hall-Dietrich Hall, Philadelphia, PA 19104, USA

A B S T R A C T

This essay discusses how incorporating qualitative analyses and insight in positivist field studies can strengthen researchers' ability to draw causal inferences. Specifically, I review how the rich institutional knowledge available in field settings can be used to increase internal validity by improving the specification of empirical models and tests and by providing greater insight into statistical results, particularly through the investigation of the causal processes linking the accounting practices and outcomes.

Introduction

Positivistic research in accounting addresses cause-and-effect questions. For example, do differences in environmental or strategic contexts *lead to* differences in management control systems? Do certain activities *drive* overhead costs? Does the adoption of a balanced scorecard system *improve* performance? However, despite this focus on causal questions, field researchers' capacity to draw strong causal inferences is hindered by their inability to conduct or study true, randomized natural experiments. Instead, researchers must rely on non- or quasi-experimental methods. The limitations in these methods give rise to concerns regarding the extent to which causation can be inferred from field-based accounting studies. The objective of this essay is to discuss how the incorporation of qualitative methods in positivistic field research can provide a powerful mechanism to enhance a study's causal inferences. In particular, researchers can take advantage of the rich institutional knowledge available in the field to strengthen the validity of the analyses through improved specification of empirical models and tests, and the ability to provide greater insight into statistical results (particularly through greater understanding of the causal processes linking the accounting practices and outcomes).

As Cook and Campbell (1979) note in their influential book on quasi-experimental field research methods, causal inferences in social science research can never be proven with certainty because the inferences depend upon many assumptions that cannot be directly verified. Any research method contains some level of uncertainty because all of the causes of observed effects or how they relate to each other are rarely if ever known. Empirical researchers must therefore attempt to assess the probability that a specific factor caused an outcome to occur. This requires choosing research methods that enhance a study's internal validity (i.e., the extent to which a study's causal conclusions are justified). As Cook and Campbell (1979, p. 11) argue, "we want to base causal inferences on procedures that reduce the uncertainty about causal connections even though uncertainty can never be reduced to zero."

Although considerable philosophical debate exists regarding the nature of causality, Mills' three well-known criteria provide a practical foundation for assessing causal relationships in empirical studies: (1) the cause has to precede the effect in time; (2) the cause and effect have to be related (i.e., co-vary); and (3) other explanations of the cause-and-effect relationship have to be eliminated. The incorporation of qualitative analyses can enhance the internal validity of positivistic field research by providing as much evidence as possible that these criteria hold in the chosen research site(s).

All empirical tests of causal relationships in non-experimental settings are susceptible to multiple threats to

* Tel.: +1 215 898 7786.
E-mail address: ittner@wharton.upenn.edu

0361-3682/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.aos.2013.10.003

validity, including correlated omitted variables and endogeneity, interactions, non-linearities, simultaneities, and measurement error, among others. To the extent possible, researchers should attempt to minimize these threats in their empirical tests. At a basic level, this requires researchers to control for any confounding influences that may affect the outcomes of interest in their regression or structural equations models. More advanced options include taking advantage of methodological improvements in statistical techniques that can enhance the ability to draw causal inferences in non-experimental studies.[1] For example, *difference-in-difference* tests can be used to compare time series changes in outcomes for groups of individuals or organizations that received a treatment (for example, implemented a balanced scorecard) relative to those that did not, with the non-treatment group serving as the control for factors other than the treatment that could influence the outcome. *Propensity scoring* methods can statistically match individuals or organizations in treatment and non-treatment groups, thereby attempting to replicate a randomized experiment as closely as possible by obtaining treatment and control groups with similar covariate distributions. *Regression discontinuity* designs can take advantage of exogenously-imposed discontinuities or cut-offs (e.g., age limits to become eligible for an incentive plan) to assign observations to treatment and non-treatment groups in the absence of random assignment. Assuming that individuals on either side of the cut-off are similar, any outcome effect should be due to the treatment. *Instrumental variables* approaches, which require the identification of instrumental variables that affect the treatment assignment but not the error term in the outcome model, can be used to control for endogeneity and other correlated omitted variables problems. Because the instrumental variable is correlated with the treatment but uncorrelated with the other determinants of the outcome, the estimated effect of the instrument on the outcome should relate only to the treatment's outcome effect, and not to the effects of variables that are correlated with the treatment.

While these and other statistical methods can improve researchers' ability to draw causal inferences, their application still requires the identification of the appropriate variables and controls to include in the tests, the timing and nature of key events (both those being studied and potentially confounding events), and the measurement of variables. Many of these issues are likely to be idiosyncratic to each research site, making them difficult to identify and incorporate in arms-length, large sample studies. In contrast, field researchers can take advantage of detailed institutional knowledge of their research context to better specify their empirical tests. More often than not, this contextual knowledge requires qualitative research to precede the specification of empirical tests. Cook and Campbell (1979, p. 93) go so far as to argue that "field experimentation should always include qualitative research to describe and illuminate the context and conditions under which research is conducted. These efforts often may uncover important site-specific threats to validity and contribute to valid explanations of experimental results in general and of perplexing or unexpected outcomes in particular."

Improved specification of empirical models and tests

One of the biggest contributions that detailed institutional knowledge can make to casual inference in field studies is identifying potentially confounding factors. Non-random assignment of individuals and organizations to treatment and non-treatment groups (for example, to adopters and non-adopters of an accounting innovation) can lead to differences in outcomes that are not due to variations in accounting practices if the two groups differ on important dimensions. In many cases, accounting and control practices are implemented concurrently with other management changes, such as the hiring of a new management team or the implementation of advanced manufacturing practices or customer satisfaction initiatives. In other cases, certain types of organizations and employees are more likely to adopt or to be assigned to a treatment than others, with these differences having direct effects on outcomes that are not due to the treatment. Factors such as economic environments, labor markets, or unionization can vary over time or across organizational units. These factors can influence the potential benefits from an accounting practice, confounding any analysis of what the practice's effects would have been in the absence of these differences. Any conclusions regarding the influence of an accounting practice on an outcome must control for confounding issues such as these.

Knowledge of the research context can help uncover the key control variables or matching criteria to include in the statistical tests in order to minimize these confounding effects. Consider, for example, Griffith and Neeley's (2009) analysis of a balanced scorecard-based pay scheme in a buildings supplies firm. One division of the firm adopted the scheme while the other did not. To improve their ability to make casual inferences, Griffith and Neeley matched branches from the two divisions (which sold similar but not identical products to the same customer base) based on postal code. Thus, the matched pairs sold similar products to the same customers and faced similar local factors such as economic cycles and labor market conditions, seemingly a nearly perfect quasi-experiment. However, their qualitative field work indicated that one division sold products that were used both inside and outside buildings, while the other focused primarily on inside usage. Thus, weather played a role in financial outcomes. In addition, one division sold goods that were mostly used for refitting buildings, while the other's goods were mostly used for new construction. Based on this firm-specific knowledge, the authors incorporated local weather conditions and different types of construction activity as additional controls in their tests.

Bol and Moers (2010) used in-depth semi-structured interviews, observations, and analysis of internal documents to better understand the introduction and diffusion of a balanced scorecard compensation plan in units belonging to a cooperative bank. Their qualitative analyses revealed a number of potential influences on scorecard

[1] See Cook and Campbell (1979) and Antonakis, Bendahan, and Lalive (2010) for discussions of these techniques.