Strategic Management Journal
Strat. Mgmt. J., **35**: 949–953 (2014)
Published online EarlyView 11 May 2014 in Wiley Online Library (wileyonlinelibrary.com) DOI: 10.1002/smj.2278
Received 11 April 2014; *Final revision received 14 April 2014*

EDITORIAL

QUANTITATIVE EMPIRICAL ANALYSIS IN STRATEGIC MANAGEMENT

RICHARD BETTIS,[1] ALFONSO GAMBARDELLA,[2] CONSTANCE HELFAT,[3]* and WILL MITCHELL[4,5]
[1] *Kenan-Flagler Business School, University of North Carolina, Chapel Hill, North Carolina, U.S.A.*
[2] *Department of Management & Technology & CRIOS, Bocconi University, Milan, Italy*
[3] *Tuck School of Business, Dartmouth College, Hanover, New Hampshire, U.S.A.*
[4] *Fuqua School of Business, Duke University, Durham, North Carolina, U.S.A.*
[5] *Rotman School of Management, University of Toronto, Toronto, Ontario, Canada*

INTRODUCTION

Strategic Management Journal seeks to encourage rigor of thought and analysis in empirical research. In this piece, we discuss quantitative empirical research. There are different approaches to rigor, and rigor does not necessarily mean complicated. Most importantly, researchers must understand their data, as well as the limitations of the empirical approaches utilized. Different empirical approaches have utility in different situations and contexts, but there is no substitute for researcher judgment. Moreover, the methods employed should neither constrain nor determine the questions asked.

THE VALUE OF GOOD DATA

Good data contain few errors, provide accurate measures of a phenomenon and underlying factors, and enable clear inferences of cause and effect using relatively simple statistical analyses. Sophisticated econometric techniques were developed in order to cope with the many problems of highly imperfect data. Collecting or obtaining access to good data solves many problems.

One of the hallmarks of strategic management research is the use of unique data; attention paid during data collection to potential pitfalls in empirical estimation can pay off later by enabling simpler or more robust analyses. This holds for archival, survey, and experimental data. Decisions regarding which archival and survey data to collect or obtain access to can benefit from an understanding of subsequent empirical issues that such data may present, as discussed below. Survey data also can suffer from common method bias, lack of longitudinal information, and retrospective and perceptual measures that appropriate survey design may help to mitigate. Randomized controlled experiments in the lab or the field have superior statistical properties, but can be limited in their external validity and their generalizability beyond the simple situations that such experiments tend to require. All of these methods have value, but require that researchers attend to their disadvantages as well as their advantages.

*Correspondence to: Constance Helfat, Tuck School of Business at Dartmouth, 100 Tuck Hall, Hanover, NH 03755, U.S.A. E-mail: constance.helfat@dartmouth.edu

THE VALUE OF FACTS

Presentation of data in well-crafted tables (e.g., Helfat, 1994; Kortum and Lerner, 1999) and in quantitative case analyses (Bresnahan and Gambardella, 1998) can very effectively shed light on many phenomena. These approaches can be used to rule out plausible alternative explanations and provide evidence much more consistent with the proposed explanation. Empirical studies need not contain statistical analysis to enable researchers to draw valid inferences.

Studies also need not necessarily seek to establish causality. Presenting facts and asking questions about possible explanations of these facts serves an important purpose. Studies that raise questions about a phenomenon can be as valuable as studies that seek to provide answers.

As part of data presentation, summary tables and basic descriptive statistics are powerful. Many, and perhaps most, primary relationships are evident in contingency tables and correlation matrices. More sophisticated analyses are often needed to deal with complicated relationships and to rule out alternatives, but the core story is often stronger and clearer if it begins with descriptive statistics and correlations.

EXPLORATORY RESEARCH

Empirical research can be purely exploratory using various graphical, algorithmic and statistical approaches. Some statistical methods such as vector autoregression are well suited to certain kinds of exploration. Such statistical methods can be used to formulate hypotheses on occasion, but if statistical methods are used, the hypotheses cannot be tested using the same data sample or one that includes the data sample as a subset.

ORIENTING PROPOSITIONS

As a mid-point between hypothesis testing and pure exploratory research, some research questions suit mid-level theoretical framing, together with subsequent analysis that first assesses the initial “orienting proposition” and then extends understanding beyond the core expectation. This may sometimes reinforce and build on the original orienting proposition, but could also overturn the original expectations. Conceptually, this approach is appropriate when existing theory provides a useful frame for a baseline argument but is not robust enough for precise hypotheses. Empirically, multiple approaches are relevant here – indeed, multi-methods might even be desirable for such studies, though not necessarily required – as they are for exploratory research.

HYPOTHESIS TESTING

Beyond the presentation of facts, good data enable the use of simpler statistical approaches, such as comparisons of means and medians and other relatively simple parametric and non-parametric statistics.

It is important to note that statistical significance is not a proof of a proposed hypothesis, but evidence in its favor. It reflects the size, nature, and composition of the sample rather than the entire population, as well as the empirical specification and variables included in the analysis. Some level of humility is always in order when characterizing the statistical significance of a variable. Significance should generally be combined with a calculation regarding the practical or economic importance of the coefficient. Significance without practical or economic importance is usually of little value. It may also be useful under some circumstances to report the power of the test so that readers can judge both Type I and Type II errors.

REGRESSION ANALYSIS

Often regression analysis is required in order to assess the role of specific variables while holding other variables constant. Here one confronts any number of general issues, including the potential for measurement error in the values of the variables, which affects both the coefficient estimates and the standard errors. A closely related problem (often termed non-classical measurement error) involves correlation of a right-hand side variable with the error term due to omitted variables.

When a researcher has panel data, fixed effects estimation can help to control for the correlation between included right-hand side variables and omitted variables in the error term. Nevertheless, fixed effects cannot be used with some estimation

DOI: 10.1002/smj

techniques (e.g., tobit), when the number of time periods is small, and when the explanatory variables of interest change slowly over time. A long-standing alternative is to include a pre-sample ("predetermined") variable to control for fixed effects; Blundell, Griffith, and Van Reenen (1995) provide an approach in which the pre-sample variable is constructed using data on the dependent variable prior to the time period of the study. Although a less good solution, a lagged dependent variable can also be used to control for fixed effects; prior year lags of the dependent variable can also be used as instruments because, as shown by Anderson and Hsiao (1982) and Arellano and Bond (1991), these variables are not correlated with the error structure. Other approaches are available as well (see Wooldridge, 2010), although some require substantial degrees of freedom.

SIMULTANEITY AND NONRANDOM SAMPLE SELECTION

Demonstrating causation in statistical analysis is easier said than done. In addition to problems in inferring causation due to measurement error and omitted variables, researchers often confront potential simultaneity of the dependent and right-hand side variables, as well as sample selection bias due to a nonrandom sample of observations. Below we outline approaches for dealing with these problems. It is important to note, however, that none of these approaches can conclusively demonstrate causality; instead, they can improve the plausibility of a causal explanation.

Logical argument based on facts

Sometimes the facts of the situation make reverse causation from right-hand side variable to the dependent variable unlikely, which can be explained through logical argument.

Rule out alternative explanations

Regression analysis can help to rule out alternative explanations, and to demonstrate correlation and statistical association between the remaining variables that is consistent with the proposed explanation.

Provide evidence regarding theoretical mechanisms

When the direction of causality is difficult to ascertain from data relating two (or more) variables, empirical examination of theoretical mechanisms that predict cause and effect can provide relevant evidence. For example, a researcher can compare outcomes in situations in which the theoretical mechanism is more versus less likely to hold, or a researcher can compare outcomes of different theoretical mechanisms to provide evidence suggestive of causality. In particular, different theories may have different predictions about the relations between two or more variables, and the empirical analysis can assess which prediction the data support.

Instrumental variables

This set of techniques can be used to deal with potentially endogenous right-hand side variables. Instruments must be strongly correlated with the right-hand side variables in question, but not with the dependent variable. Such instruments can be hard to find, and a bad instrument is worse than no instrument. (For recent discussion of this issue in management research, see Bascle, 2008; Semadeni, Withers, and Certo, 2014.) Recently, tests for weak instruments have been devised. However, while indicating whether the chosen instruments may have too low a correlation with the endogenous variable, such tests do not establish that: (a) the instrument does not affect the dependent variable directly other than only through the endogenous variable; (b) the instrument is exogenous; (c) the instrument has a logical relationship with the endogenous variable. In other words, although these tests can identify a weak instrument, the opposite does not imply (a), (b), or (c). Instead, a researcher must provide good arguments that an instrument has a logical relationship with the endogenous variable, is correlated with the dependent variable only through the endogenous variable, and is not itself endogenous.

Sample selection correction

This set of techniques can help to control for bias due to a nonrandom sample, generally using two-stage estimation. The first stage equation requires inclusion of one or more variables correlated with the variable that is subject to selection

DOI: 10.1002/smj

bias in the second stage (the dependent variable in the first stage), but uncorrelated with the dependent variable in the second stage. Such variables can be hard to find, as in the case of instrumental variables more generally. Also, unlike in standard instrumental variable regressions, a sample selection model that employs a non-linear correction factor in the second stage (typically the inverse Mills ratio) can converge even in the absence of sufficient or accurate identification. However, we cannot attribute our results to the selection effect, as they could be produced by the nonlinear impacts on the dependent variable. As for instrumental variables estimation more generally, to correctly identify selection effects, there is no substitute for logical and convincing arguments based on knowledge of the facts and the problem at hand.

Matching techniques

Matched samples have been used occasionally in strategic management, particularly in cases with relatively few observations of the phenomenon of interest (e.g., Chatterjee, 1986). In observational studies, a researcher can match a sample of firms (the treatment group) that exhibit a phenomenon of interest with a set of control firms (the control group) that have similar characteristics but do not exhibit the phenomenon of interest, and compare outcomes. Matching techniques provide another approach for dealing with sample selection bias regarding which firms exhibit the phenomenon under investigation. Sophisticated matching techniques–which work best for, and often require, larger sample sizes–include *propensity score matching* and *coarsened exact matching*. These techniques also have limitations, including with respect to the criteria for matching and sample size. For example, the techniques can only match on, and therefore control for, observable variables. The techniques often need large sample sizes to implement them correctly as well. (For a review, see Stuart, 2010.)

Differences-in-differences

Quasi-natural experiments in which an exogenous shock affects one group (the treatment group) but not another (the control group) can help to isolate causality. However, it can be difficult to find a control group that is free of sample selection bias. In addition, estimation can be subject to reversion to the mean. Differences-in-differences estimation can be used in lab and field experiments that may be able to avoid sample selection problems, but at the cost of more limited generalizability.

Granger causality

This technique does *not* show that the lagged variables on the right-hand side cause the dependent variable; instead, the technique can establish that one set of variables precedes another variable in time.

NON-RESULTS

Empirical analysis may produce a statistically insignificant coefficient on a variable of interest—often termed a "non-result." Lack of results can be of substantial interest. First, such a finding is relevant in studies that seek to replicate the results of prior studies. Secondly, the lack of results predicted by theory can lead to a reevaluation of the theory. Non-results, however, can also come from problems of statistical estimation, inappropriate application of the theory, or poor data. Thus, it is important to demonstrate that a non-result holds using multiple approaches, to conduct appropriate statistical analyses, and to rule out obvious alternative reasons for the non-result other than that the hypothesized theory may not hold.

DATA SNOOPING AND HYPOTHESIS TESTING

The practice known as "data snooping" or "searching for asterisks" consists of statistically examining a database to find models that include statistically significant variables. Hypotheses are then formulated to explain the significant variables. Finally, results using the same data, perhaps supplemented with additional data, are presented as theory-driven hypothesis testing. (For recent discussion of this issue in strategic management, see Bettis, 2012). In a variant of this practice known as "p-hacking", researchers adapt protocols for experiments in order to produce statistically significant results. These practices are totally inconsistent with the proper use of statistical hypothesis testing and result in the reporting of inappropriate and inflated *p*-values supporting contrived hypotheses. Such practices, which clearly violate professional norms, should always be avoided.

CONCLUSION

There are many valid approaches to quantitative empirical research. The use of a varied toolkit, by researchers cognizant of both the advantages and pitfalls of each approach, advances strategic management research. In this regard, the research question and data should drive the methods, rather than methods driving the research. Claims regarding the results also must be appropriately calibrated to what the data and methods allow.

REFERENCES

Anderson TW, Hsiao C. 1982. Formulation and estimation of dynamic models using panel data. *Journal of Econometrics* **18**(1): 47–82.

Arellano M, Bond S. 1991. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies* **58**(2): 277–297.

Bascle G. 2008. Controlling for endogeneity with instrumental variables in strategic management research. *Strategic Organization* **6**(3): 285–327.

Bettis RA. 2012. The search for asterisks: compromised statistical tests and flawed theories. *Strategic Management Journal* **33**(1): 108–113.

Blundell R, Griffith R, Van Reenen J. 1995. Dynamic count data models of technological innovation. *Economic Journal* **105**: 333–344.

Bresnahan T, Gambardella A. 1998. The division of inventive labor and the extent of the market. In *General Purpose Technologies and Economic Growth*, Helpman E (ed). The MIT Press: Cambridge, MA; 253–281.

Chatterjee S. 1986. Types of synergy and economic value: the impact of acquisitions on merging and rival firms. *Strategic Management Journal* **7**: 119–139.

Helfat C. 1994. Firm-specificity in corporate applied R&D. *Organization Science* **5**(2): 173–184.

Kortum S, Lerner J. 1999. What is behind the recent surge in patenting? *Research Policy* **28**: 1–22.

Semadeni M, Withers MC, Certo ST. 2014. The perils of endogeneity and instrumental variables in strategy research: understanding through simulations. *Strategic Management Journal* (Forthcoming) (published online 25 June 2013).

Stuart EA. 2010. Matching methods for causal inference: a review and a look forward. *Statistical Science* **25**(1): 1–21.

Wooldridge JM. 2010. *Econometric Analysis of Cross Section and Panel Data* (2nd edn). The MIT Press: Cambridge, MA.

DOI: 10.1002/smj