Strategic Management Journal
Strat. Mgmt. J., **24**: 1231–1237 (2003)
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/smj.352

RESEARCH NOTES AND COMMENTARIES

DATA ANALYTIC TRENDS AND TRAINING IN STRATEGIC MANAGEMENT

CHRISTOPHER L. SHOOK,[1]* DAVID J. KETCHEN JR,[2] CYNTHIA S. CYCYOTA[3] and DILENE CROCKETT[4]
[1] *College of Business, Auburn University, Auburn, Alabama, U.S.A.*
[2] *College of Business, Florida State University, Tallahassee, Florida, U.S.A.*
[3] *U.S. Air Force Academy, USAF Academy, Colorado, U.S.A.*
[4] *College of Business Administration, University of Texas at Arlington, Arlington, Texas, U.S.A.*

Data analysis is a key element of the research process. Accordingly, appropriate doctoral training in data analysis is vital to the strategic management field's future. We used a two-study design to evaluate quantitative data analysis trends and doctoral training. An analysis of Strategic Management Journal *articles from 1980 to 2001 reveals that, contrary to some predictions, the use of general linear model techniques such as regression has increased over time. However, the use of more specialized techniques, including those suitable for examining longitudinal data, discrete events, and causal structure, has also grown substantially. A survey of recent doctoral graduates shows that, although skilled with general linear models, many are ill prepared to use specialized techniques. Based on our findings, we offer suggestions aimed at bridging gaps between what doctoral students (and other researchers) know and what they need to know about data analysis.* Copyright © 2003 John Wiley & Sons, Ltd.

Young scholars represent the future of our field. Given this importance, it is surprising that little attention has been devoted to understanding whether or not doctoral students are being provided adequate methodological training to advance the field. The last thorough discussion of doctoral training is over a decade old (i.e., Summer *et al.*, 1990), despite seemingly important changes in the interim. Indeed, Hitt, Gimeno, and Hoskisson (1998: 17) assert that strategic management research is moving 'beyond cross-sectional, multiple regression approaches to methods more attuned to the specific problems and issues likely to influence strategy research' such as network analysis, event studies, and Poisson/negative binomial regression. If this statement is true, doctoral students should be trained in certain methods that are not yet prominent but that offer promise for advancing strategic management research.

This research note presents two studies that each tackle an important issue. First, we test Hitt *et al.*'s assertion by tracking trends in the use of data analytic techniques. Understanding what methods are increasing and decreasing in use has important implications for the training of doctoral students. Second, we seek to understand the level of mastery recent doctoral graduates possess with both traditional and specialized methods. By revealing

Key words: strategy research; statistical methods; research trends; doctoral training
*Correspondence to: Christopher L. Shook, College of Business, Auburn University, 401 Lowder Business Building, Auburn, AL 36849-5241, U.S.A.

Received 12 September 2000
Final revision received 29 April 2003

gaps between what doctoral students know and what they need to know about quantitative methodology, we identify areas that deserve additional emphasis in doctoral training. Indeed, leading doctoral programs to steer students' research methods training toward areas of need is our overall objective.

STUDY 1: DATA ANALYTIC METHODS IN *SMJ*

Sample and data

Our sample was drawn from articles published in the *Strategic Management Journal* (*SMJ*) from 1980 to 2001. We randomly selected half of the articles published each year ($n = 523$). Of these articles, 297 presented original empirical studies. We coded only the data analytic procedures used to test studies' hypotheses or to answer research questions. For example, Subramaniam and Venkatraman (2001) first factor analyzed data to assess the validity of survey measures, examined correlations among the items, and, finally, used hierarchical regression to test hypotheses. Only hierarchical regression was coded because it was the method that directly examined the hypotheses.

To check for coding consistency, both coders coded studies from four of the years. The coders' percentage agreement (Zegers, 1991) was 81 percent, which is comparable to similar studies (e.g., 83% = Ford, MacCallum, and Tait, 1986). Following discussion and resolution of coding differences, the coders coded alternating years.

Analysis and results

We computed periodic percentage use indices (PUIs) for each method (Stone-Romero, Weaver, and Glenar, 1995). Specifically, we divided the frequency of use for each method by the number of articles coded per period. Thus, a PUI of 0.30 indicates that 30 percent of the studies during a time period relied on a focal method. To assess trends in use of methods over time, we computed correlations between years and annual PUI values for each method (Stone-Romero *et al.*, 1995). The correlations are presented in Table 1 along with PUI values by decade (1980s, 1990s, and 2000s). Although the grouping by decades involves an unequal number of years (i.e., only 2 years in the 2000s), this grouping is for descriptive purposes only and does not impact the correlations. To facilitate comparison to Hitt *et al.*'s (1998) predictions about the increasing importance of certain methods, we grouped methods based on their categorizations, and present PUIs for the categories as well as for the individual techniques.

As noted above, Hitt *et al.* (1998) assert that strategy research is moving beyond traditional tools such as multiple regression. Years earlier, Camerer and Fahey (1988) exhorted the field to escape the domination of the 'regression paradigm.' Some basic techniques have indeed fallen out of favor. For example, the PUI for simple tests of mean differences was 0.30 in the 1980s, but only 0.07 in the 1990s. This technique was absent from studies in 2000–01.

However, our results show that general linear models (GLM) remain the dominant group of techniques. Indeed, in contrast to Hitt *et al.*'s assertion and Camerer and Fahey's plea, GLM use has increased over time ($r = 0.499$, $p < 0.05$). Collectively, GLM techniques were used in 57 percent of studies in the 1990s and 63 percent of the studies in 2000–01. Multiple regression and hierarchical regression are by far the most popular GLM techniques, accounting for 45 percent of all studies in the 1990s and 56 percent in 2000–01. The use of these two techniques is highly correlated with the passage of time (0.567 and 0.681, respectively), indicating that their use has risen significantly since the 1980s.

Meanwhile, some more specialized methods have gained adherents, in partial support of Hitt *et al.*'s (1998) characterization of the field. The use of methods to analyze longitudinal data, methods appropriate for analyzing discrete events, and methods appropriate for discovering causal structure all are significantly and positively correlated with the passage of time (0.422, 0.549, and 0.537 respectively). Perhaps logistic regression and structural equation modeling offer the most dramatic increases. These techniques were little used in the 1980s but have substantial PUIs in the other periods (0.12 and 0.11 for logistic regression; 0.06 and 0.13 for structural equation modeling). Overall, specialized techniques are growing in use, but this growth is not coming at the expense of GLM techniques.

Table 1. Summary of data analytic technique usage

Data analytic technique	1980s	1990s	2000s	Correlation with year
Frequencies	0.03	0.02	0.00	−0.229
Nonparametric tests	0.13	0.02	0.00	−0.376†
Correlations	0.03	0.01	0.00	−0.294
Tests of mean differences	0.30	0.07	0.00	−0.704**
t-Tests	0.27	0.07	0.00	−0.622**
Other tests of means	0.03	0.00	0.00	−0.308
General linear models	0.42	0.57	0.63	0.499*
ANOVA	0.14	0.08	0.07	−0.248
ANCOVA	0.01	0.02	0.00	−0.085
MANOVA	0.01	0.01	0.00	0.204
MANCOVA	0.02	0.01	0.00	−0.227
Simple regression	0.03	0.00	0.00	−0.358
Multiple regression	0.19	0.34	0.39	0.567**
Hierarchical regression	0.02	0.11	0.17	0.681***
Longitudinal data methods	0.01	0.02	0.04	0.422*
Fixed effects models	0.00	0.00	0.00	0.292
General time series analysis	0.00	0.01	0.00	0.052
Pooled time series analysis	0.00	0.00	0.02	0.361†
Variance decomposition	0.01	0.01	0.02	0.273
Explicitly dynamic methods				
Event history/hazard studies	0.00	0.03	0.02	0.413†
Discrete events methods	0.04	0.16	0.16	0.549**
Discriminant analysis	0.04	0.01	0.00	−0.369†
Financial event study	0.00	0.03	0.05	0.453*
Logistic regression	0.00	0.12	0.11	0.757***
Methods for analysis of interdependence among firms				
Network analysis	0.00	0.01	0.00	0.258
Methods explicitly accounting for firm heterogeneity				
Cluster analysis	0.02	0.00	0.00	−0.083
Causal structure methods	0.02	0.08	0.13	0.537**
Path analysis	0.01	0.02	0.00	−0.209
Simultaneous equations	0.01	0.00	0.00	−0.258
Structural equation modeling	0.00	0.06	0.13	0.732***
Methods to analyze decision making				
Causal mapping	0.00	0.01	0.02	0.349

† $p < 0.10$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

STUDY 2: PhD TRAINING IN DATA ANALYTIC METHODS

Sample and data

The data for this study were drawn from 77 strategic management researchers who had attended the Academy of Management's Business Policy and Strategy Division Doctoral Consortium between 1996 and 2001. The consortium is for Ph.D. students who have defended their dissertation proposals and have been nominated by their institution as its 'best' eligible student.

A total of 291 people attended the consortia. We were not able to locate contact information for 41. An initial request to complete a survey and a follow-up request were e-mailed to the remaining 250 scholars along with a link to a web page. On the web page, respondents were asked to indicate their proficiency level with various traditional methods and a list of specialized methods adapted from Hitt *et al.* (1998). More specifically, we asked the respondents 'When you left graduate school, how competent were you with each method?' as well as 'To what extent are you competent now with these methods?' Anchors for the scales were 1 ('not at all') and 5 ('to a great extent').

Two scholars responded that they were unwilling/unable to complete the instrument. Usable responses were collected from 77 scholars for a response rate of 31 percent. Responses were received from 11 people who attended the consortium in 1996, nine in 1997, 18 in 1998, nine in 1999, 14 in 2000, and 16 in 2001. The extrapolation procedure suggested by Armstrong and Overton (1977) was used to assess potential response bias. We compared the first 25 percent and the last 25 percent of the respondents on the survey items and on demographic variables (gender, doctoral training inside or outside the United States, and year of consortium attendance). No significant differences were found; thus, there was no evidence of a response bias.

Analysis and results

Table 2 shows the means and standard deviations for the survey items. Traditional techniques such as correlations, tests of mean differences, and general linear models all have mean responses of 3.3 or higher on a five-point scale at the time of degree completion. Four techniques out of seven have means well above four, including simple and multiple regression (4.69 and 4.61, respectively). Thus, survey respondents believe they were well trained in graduate school with these traditional techniques.

Table 2 also reveals that many doctoral students are graduating with little confidence in their ability to use the methods Hitt *et al.* (1998) predict will be more important in the future. Only logistic regression (3.48) and exploratory factor analysis (3.69) had means above the mid-point of the scale. Of the remaining 21 methods, eight had means between two and three, while 13 had means below two. In most cases, the standard deviations were above one. This suggests that although the typical researcher was not adept with these methods at graduation, some researchers were. One possibility was that a small group of researchers were proficient with many of the methods. Instead, we found that most new scholars are specializing in a few techniques. Fifty-four respondents (70%) were skilled (i.e., rated as a four or five) with five or fewer of the 23 emerging techniques. Only eight respondents were skilled with more than 10 techniques and none exceeded 17.

We were interested in whether respondents' competencies were significantly higher with traditional techniques than with specialized techniques. Because multiple regression is the most widely used traditional technique, we compared competence with it vs. the specialized methods via paired-samples t-tests using the Bonferroni adjustment for multiple comparisons. Competence with multiple regression was significantly higher ($p < 0.01$) than competence with all 23 specialized methods both at completion of the Ph.D. and at present.

We next assessed if the respondents' perceived competencies with each specialized method changed between graduation and the present. Using paired-sample t-tests with the Bonferroni adjustment, there was little significant change uncovered. Competency with only one technique (i.e., network analysis) out of 23 showed a significant increase (2.11 to 2.42; $p < 0.05$). These results suggest that recently matriculated scholars have only maintained their level of familiarity with specialized techniques since graduation, not improved it.

Finally, we were interested in whether training with specialized methods appears to be improving over time. We first computed correlations between competence at graduation with each method and the year the respondent's Ph.D. was earned. Of the 23 methods, competence with network analysis, self-selection models, and simultaneous models was positively and significantly correlated with the year of the Ph.D. ($r = 0.27$, $p < 0.05$; $r = 0.34$, $p < 0.01$; and, $r = 0.24$, $p < 0.05$, respectively). Thus, more recent graduates left graduate school possessing more confidence about these three techniques than did earlier graduates. Next, we examined the relations between reported present competencies and the year the respondent's Ph.D. was earned. None of these correlations were significant.

DISCUSSION

The results of our studies should be viewed in light of their limitations. In Study 1, the 297 articles examined were all published in *SMJ*. This sample appears appropriate because *SMJ* has been shown to be the premier outlet for strategy research (Tahai and Meyer, 1999). It is possible, however, that other journals might feature a different mix of methods. In Study 2, we surveyed

Table 2. Summary of competence levels with quantitative data analysis techniques

	Upon completion of Ph.D.		At present	
	Mean	S.D.	Mean	S.D.
Traditional techniques				
Correlations	4.49	0.74	4.58	0.69
Tests of mean differences				
t-Tests	4.39	0.76	4.42	0.72
General linear models				
ANOVA and ANCOVA	3.78	1.04	3.90	1.10
MANOVA and MANCOVA	3.34	1.15	3.46	1.17
Simple regression	4.69	0.61	4.69	0.69
Multiple regression	4.61	0.78	4.67	0.65
Hierarchical regression	3.81	1.22	3.97	1.38
Specialized techniques				
Longitudinal data methods				
Panel data analysis	2.88	1.46	2.85	1.58
Repeated measures analysis	2.15	1.10	2.03	1.17
Explicitly dynamic methods				
Event history	2.50	1.22	2.53	1.28
Partial adjustment models	1.45	0.73	1.45	0.76
Dynamic adjustment models	1.38	0.63	1.43	0.76
Discrete events methods				
Financial event study	2.23	1.39	2.27	1.39
Logistic regression	3.48	1.24	3.68	1.22
Multinomial logistic regression	2.80	1.27	2.95	1.36
Poisson/negative binomial regression	2.39	1.39	2.41	1.42
Methods for analysis of interdependence among firms				
Diffusion models	1.59	0.91	1.69	0.90
Network analysis	2.11	1.21	2.42	1.26
Multidimensional scaling	2.19	1.06	2.15	1.27
Causal structure methods				
Path analysis	2.78	1.30	2.79	1.39
Sample selection models	2.45	1.39	2.25	1.45
Seemingly unrelated regression	1.74	1.12	2.03	1.25
Self-selection models	1.77	1.22	1.72	1.18
Simultaneous equations	2.49	1.28	2.60	1.38
Structural equation modeling—structural models	2.81	1.32	2.88	1.39
Methods to account for imperfect measurement of constructs				
Exploratory factor analysis	3.69	1.31	3.72	1.25
Confirmatory factor analysis	2.94	1.41	3.09	1.43
Methods to analyze decision making				
Repertory grid	1.38	0.76	1.43	0.82
Cognitive mapping	1.66	0.92	1.71	1.02
Policy capturing	1.49	0.90	1.65	1.10

recent graduates who had attended doctoral consortia. Other graduates might have different levels of method competence than our respondents. However, because admission to the consortia is competitive, our respondents should be among the better-trained students. As such, our finding that students lack competence in many emerging methods may actually understate the situation *vis à vis* recent graduates in general. Also, our focus on quantitative data analysis led us to exclude qualitative research in both studies. This decision simply reflects the intent of the paper and is not meant to imply that qualitative techniques are not valuable. Despite these limitations, our findings offer important implications for both the research process in general and for doctoral training.

Effective knowledge development through research depends on a match between research methodology and the phenomenon under investigation. In the case of strategic management, the relationships of interest are often complex and multifaceted. This led Camerer and Fahey (1988: 449) to criticize the field's reliance on regression techniques because 'regression techniques are notoriously weak at establishing causation and disequilibrium effects ... are inexorably lumped together, in residuals, with other kinds of unexplained variance.' Study 1 revealed that the use of more specialized techniques is in fact on the rise, as asserted by Hitt *et al.* (1998) and encouraged by Camerer and Fahey (1988). However, the rise of these methods is not coming at the expense of regression. Instead, the reliance on regression has grown since the 1980s (along with the use of more specialized methods) and has been offset by the decline of very basic techniques such as t-tests and other tests of mean differences.

What are the implications of these trends for knowledge development? Examining the fit between studies' methodology and phenomena of interest was beyond the scope of this investigation. Perhaps some studies were limited to general linear models because the data were cross-sectional. However, it seems unlikely that general linear models offered the best match for the data in more than half of the studies since 1990. Nonetheless, such techniques were used in 57 percent of studies in the 1990s and in 63 percent in 2000–01. Hitt *et al.* (1998: 17) note that regression is 'ineffective for testing hypotheses in data that is nonexperimental and laden with nonrecursive relationships.' Yet, few strategy studies use experimental data, and many conceptual models involve feedback loops. Thus, one reasonable conclusion is that many researchers are not fully exploiting their data. As Bergh (1993) demonstrated, the use of suboptimal techniques can lead to erroneous conclusions, thereby inhibiting knowledge development. Conversely, the increased use of more specialized techniques whose capabilities and assumptions better reflect the nature of strategic phenomena is encouraging *vis à vis* knowledge development. Thus, the findings of Study 1 offer cause for both concern and encouragement.

The findings of Study 2 were perhaps more alarming. A core set of techniques with which recent Ph.D.s are competent emerged from our study. However, this core appears relatively small and does not include the techniques predicted by Hitt *et al.* (1998) to be vital to current and future knowledge development. While not every doctoral student needs to master every technique, our results highlight that doctoral training may not be keeping pace with data analytic trends and future research needs. The potential implications are dramatic. To the extent that a strategy scholar lacks sufficient methods training, his/her ability to contribute to the body of knowledge is inhibited. Most scholars may limit themselves to research opportunities that fit with their narrow skill set. In this sense, the entire discipline suffers as a result of inadequate doctoral training.

Accordingly, doctoral programs should work to close the gap between what their students know about methods and what they need to know. Given the stakes, a variety of innovative actions may be warranted. Quantitative methodology seminars within a management or strategic management department could be team taught, with a series of specialized techniques each covered by the professor most skilled in the method. With a warning to keep in mind differences in disciplinary assumptions (Hirsch, Friedman, and Koza, 1990), strategy students could be directed to other departments to gain skills in certain methods that are already popular in those areas. For example, most of our respondents possess little confidence in their ability to use event studies and methods for understanding decision making (e.g., repertory grid technique and cognitive mapping). Students could learn about the former through finance departments and the latter through psychology departments. Looking beyond the boundaries of individual institutions, cooperation among universities in large metropolitan areas could result in joint methodology classes, while distance learning could aid students in other areas. Overall, faculty should think creatively to address doctoral students' apparent lack of skills with important, emerging techniques.

We also found that few new scholars have developed additional data analytic competencies in the years immediately following graduation. These results highlight the need for faculty to renew and enhance their skills. Many management departments host visiting scholars for a few days each semester. In our experience, visitors are generally selected based on their success in developing ideas and theories. Given our findings, perhaps some of these guests should be chosen based on their methodological expertise. Professional meetings offer another potential forum for enhancement. Currently, few sessions at prominent academic conferences are focused on how to use specific methods. We suggest that an increase in the amount of workshops devoted to specialized methods could serve as a valuable mechanism for skill development.

CONCLUSION

In discussing doctoral education over a decade ago, Summer *et al.* (1990) noted that the openness and creativity of the strategic management field requires new scholars to exit graduate school with an understanding of a variety of research techniques. Our results suggest that although methods capable of capturing complex strategic phenomena are becoming more popular in the literature, many recent graduates lack the methodological diversity called for by Summer *et al.* (1990). It thus appears that there is a certain 'madness to our methods' that the field in general and doctoral educators in particular must resolve.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the insights offered by Jim Combs, G. Tomas Hult, Chuck Kacmar, Micki Kacmar, and Chuck Snow.

REFERENCES

Armstrong JS, Overton TS. 1977. Estimating nonresponse bias in mail surveys. *Journal of Marketing Research* **14**: 396–402.

Bergh D. 1993. Don't 'waste' your time! The effects of time series errors in management research: the case of ownership concentration and research and development spending. *Journal of Management* **19**(4): 897–914.

Camerer C, Fahey L. 1988. The regression paradigm in strategy research: a critical appraisal and suggested directions. In *Strategic Management Frontiers*, Vol. 10, Grant J (ed). JAI Press: Greenwich, CT; 443–459.

Ford JK, MacCallum RC, Tait M. 1986. The application of exploratory factor analysis in applied psychology: a critical review and analysis. *Personnel Psychology* **39**: 291–314.

Hirsch PM, Friedman R, Koza MP. 1990. Collaboration or paradigm shift? Caveat emptor and the risk of romance with economic models for strategy and policy research. *Organization Science* **1**: 87–97.

Hitt MA, Gimeno J, Hoskisson RE. 1998. Current and future research methods in strategic management. *Organizational Research Methods* **1**: 6–44.

Stone-Romero EF, Weaver AE, Glenar JL. 1995. Trends in research design and data analytic strategies in organizational research. *Journal of Management* **21**(1): 141–157.

Subramaniam M, Venkatraman N. 2001. Determinants of transnational new product development capability: testing the influence of transferring and deploying tacit overseas knowledge. *Strategic Management Journal* **22**(4): 359–378.

Summer CE, Bettis RA, Duhaime IH, Grant JH, Hambrick DC, Snow CC, Zeithaml CP. 1990. Doctoral education in the field of business policy and strategy. *Journal of Management* **16**: 361–398.

Tahai A, Meyer MJ. 1999. A revealed preference study of management journals' direct influences. *Strategic Management Journal* **20**(3): 279–296.

Zegers FE. 1991. Coefficients for interrater agreement. *Applied Psychological Measurement* **15**(4): 321–333.