MATCHING METHOD TO PARADIGM IN STRATEGY RESEARCH: LIMITATIONS OF CROSS-SECTIONAL ANALYSIS AND SOME METHODOLOGICAL ALTERNATIVES

HARRY P. BOWEN[1] and MARGARETHE F. WIERSEMA[2]
[1]*Department of Economics, University of California, Irvine, California, U.S.A.*
[2]*Graduate School of Management, University of California, Irvine, California, U.S.A.*

A central focus of empirical research in strategic management has been to understand the relationships associated with the structure–strategy–performance paradigm. To examine these relationships, investigators have relied extensively on cross-sectional methods that embody the implicit assumption that model parameters are stable across firms and over time. Yet, many of the theoretical constructs used in strategic management have clear firm- and time-specific components. Hence, it might be expected that the parameters of the relationships investigated empirically will vary across firms and over time. Whereas recent research has raised concerns about the use of cross-sectional analysis when parameters vary over time, little attention has been given to the issue of parameter variability across firms. Given the focus of strategy researchers on firm-level effects and the predominant reliance on cross-sectional analysis, accounting for across-firm variability is a significant methodological issue. Failure to account for such variability can lead to biased parameter estimates and incorrect inferences. This paper argues for the adoption of alternative methods that can overcome the limitations of a cross-sectional analysis and it offers guidance on how researchers can proceed to use these alternative methods to explicitly incorporate or test for variation in model parameters across firms or over time. Copyright © 1999 John Wiley & Sons, Ltd.

INTRODUCTION

A central focus of empirical research in strategic management has been to understand the structure–strategy–performance paradigm. In this regard, researchers have investigated a variety of hypotheses concerning the link between a firm's competitive environment, its strategy and its performance, and similarly, how managers' strategic choices influence firm performance. Most of the theoretical constructs that underlie these investigations have clear firm-specific and time-related

aspects. In particular, a central premise of most strategy research is that of interfirm variation— that firms do differ in their response to environmental factors and in their determination of firm strategy and performance results. Therefore, the links between strategic choice and firm performance, or the firm's strategic responses to changes in its competitive environment, are all expected to depend on firm-specific characteristics. Despite this emphasis on firm-specific effects, strategy researchers have used mostly cross-sectional methods to empirically investigate the relationships of interest. But such methods implicitly assume that model parameters are stable (constant) across firms and over time. The use of cross-sectional methods is therefore glaringly at odds with the firm-specific aspects of the theoretical models employed in strategic management. Moreover, if cross-sectional methods are used to

estimate models whose parameters do in fact vary across firms (or over time) the resulting estimation may fail to yield statistically valid inferences. Not only can the estimation fail to identify the true model parameters but the estimates that are obtained may not be efficient.[1]

Despite its obvious importance, the issue of parameter variation and the associated limitations of cross-sectional methods have received only limited attention in the empirical strategy literature. When the issue has been discussed, the focus has mostly been the issue of time-varying parameters, and thus that inferences may be dependent on the sample year chosen for analysis.[2] For example, Bergh (1995: 1697) states in his review of empirical research on diversification and performance that 'researchers have not included time-related change in their empirical models, either as a structural component or as a factor.' Similarly, Lubatkin and Chatterjee (1991) note that the reliance on cross-sectional data for a single year prevents strategy researchers from accounting for any trend effects in their investigation. Rumelt (1991) cautions researchers that a reliance on a single year of data can result in a failure to detect the true nature, if any, of the relationship investigated. In yet another study, Hill and Hansen (1991: 187) state that 'cross-sectional studies suffer from an inability to determine the true causal relationship.' Despite such cautions, Bergh and Holbein (1997) note that most researchers still fail to examine the stability of their empirical relationships over time despite the increasing use of longitudinal studies in strategy research.

Whereas concerns have been raised about the issue of parameter instability over time, the seemingly more important issue of across-firm variation has gone almost unnoticed. But across-firm variation appears particularly important for strategy research since the hypothesized linkages between strategic choice and firm performance, or the firm's strategic response to changes in its competitive environment, can all be expected to depend on characteristics specific to each firm. Therefore, the possibility of parameter variability across firms would seem a significant methodological issue in empirical strategy research.

Despite recent cautions, cross-sectional methods remain the predominate mode of analysis in empirical strategy research. This fact is underscored by a review of the methodologies used by empirical studies published in the *Strategic Management Journal* between 1993 and 1996. Among the 90 studies surveyed, regression-based techniques such as linear regression, ANOVA, MANOVA, logit, and discriminant analysis were predominate and 75% of the studies used only a single year of data.[3] Whereas a reliance on data for only a single year could be indicative of data limitations, over 70 percent of the studies surveyed used secondary data sources (e.g., Compustat) which often provide firm-level data over a number of years. The simple fact is that few studies have attempted to exploit these larger data samples by using either time-series or pooled time-series, cross-sectional data methods. Overall, the survey results validate the concerns of Bergh (1995), Hill and Hansen (1991), and Lubatkin and Chatterjee (1991) regarding the widespread use of single-year, cross-sectional data samples in strategic management research.

As to the issue of firm- or time-specific effects, only two of the studies surveyed attempted to control for firm-specific effects and only six attempted to control for time-specific variation. Moreover, in those cases where pooled data sets were used, most studies simply assumed parameter stability across firms or over time—no formal test of this assumption was made. Of equal concern is that almost 90 percent of the studies using multivariate techniques failed to test for the presence of heteroscedasticity. This is

---

[1] While it is common to regard unbiasedness as a necessary property for an estimator this is not necessarily true. In fact, a large body of econometrics literature shows that by a squared error loss criterion there is always an estimator that dominates the (unbiased) least-squares estimator (original reference is James and Stein, 1961). Under the mean square error loss criterion, specification errors such as those implicitly discussed in the paper involve a trade-off between bias and efficiency. Efficiency therefore remains important for determining the 'best' estimator.

[2] While this issue has recently received attention, methodological concerns are not new in strategic management research. For example, issues of measurement continue to be identified as problematic (Chakravarthy, 1986; Godfrey and Hill, 1995; Robins and Wiersema, 1995; Venkatraman and Grant, 1986) while others have evaluated the application of alternative statistical techniques to major research issues in the field (Ketchen and Shook, 1996).

[3] A detailed survey of the 90 *SMJ* studies tabulated a variety of attributes including statistical method, source of data, time period studied, controlling for parameter constancy across firms and over time, and testing for heteroscedasticity. This survey is available from the authors upon request.

troubling not only because heteroscedasticity commonly arises in cross-sectional data but, as discussed below, heteroscedasticity can be expected to arise when cross-sectional methods are used to estimate a model whose parameters do vary across firms or over time. Failure to correct for heteroscedasticity when it exists leads to parameter estimates that are not efficient and hence incorrect inferences regarding the statistical significance of a model's independent variables.

As the survey results indicate, empirical research in strategic management has often failed to give appropriate attention to some common statistical issues that can arise in a cross-sectional framework and it has generally failed to deal with the important issue of parameter variation across firms or over time. Such failings may reflect that, despite the continued reliance on cross-sectional methods, a discussion of the assumptions and limitations of cross-sectional methods has been absent from the strategic management literature.

The purpose of this paper is therefore twofold: first is to discuss and clarify the limitations of a cross-sectional analysis; second is to emphasize the importance of across-firm variation for strategy research—why it can be expected to arise and why the continued reliance on cross-sectional methods is then inappropriate. In this context, the paper argues for the adoption of alternative methods that can overcome the limitations of a cross-sectional analysis and it offers guidance on how researchers can proceed to use these alternative methods to explicitly incorporate or test for variation in model parameters across firms or over time.

REGRESSION ANALYSIS IN STRATEGY RESEARCH

The continuing reliance of strategic management research on cross-sectional techniques warrants a discussion of what a cross-sectional analysis specifically investigates and what assumptions are embodied in this approach. This section addresses this issue in the context of cross-sectional regression methodologies and focuses on the assumptions these methods impose regarding coefficient variation, and the statistical issues that arise when these assumptions are violated.

Overview of methodological issues

Consider the following simple yet general representation of a relationship hypothesized to exist between two firm-level variables Y and X (e.g., performance and diversification):

Model 1 $\quad Y_{it} = \alpha_{it} + \beta_{it}X_{it} + u_{it}$

where the disturbance u_{it} is assumed to be normally distributed with mean zero and constant variance. As written, the coefficients in Model 1 have both firm (i) and time (t) subscripts and thus vary across firms and over time. Using data for a single year, a cross-sectional analysis would estimate relationship in Model 1 using a model of the form:

Model 2 $\quad Y_{it} = \alpha_t + \beta_t X_{it} + u_{it}$

In comparison to Model 1, Model 2 assumes that the coefficients at time t are constant across firms (i.e., $\alpha_{it} = \alpha_t$ and $\beta_{it} = \beta_t$). This assumption is needed since otherwise the number of coefficients to estimate would exceed the number of observations. But there is a second (implicit) assumption associated with Model 2: the coefficients are also constant over time (i.e., $\alpha_t = \alpha$ and $\beta_t = \beta$). This latter assumption is made when inferences based on the results from estimating Model 2 are thought to apply to years different from the sample year. When these assumptions about parameter constancy are valid, cross-sectional estimation of Model 2 can be thought to be uncovering a stable, long-run relationship between Y and X. But when these assumptions are not valid Model 2 is incorrectly specified and estimation using cross-sectional data can fail to identify the true model parameters and lead to incorrect inferences regarding parameter significance.[4]

Using cross-sectional analysis in the presence of parameter variation

The issues that arise when cross-sectional methods are used to estimate a model whose parameters vary across firms or over time can be illustrated by assuming that the parameters α_{it}

---

[4] The issue of parameter stability also has implications for forecasting. See Pant and Starbuck (1990) for discussion of some of the pitfalls associated with forecasting techniques in management research.

and β_{it} each depend on some variable that varies across firms and over time. Specifically, let the values of these parameters be determined according to the following relationships:

Model 3 $\alpha_{it} = a_0 + a_1 W_{it} + v_{it}$

Model 4 $\beta_{it} = b_0 + b_1 Z_{it} + \eta_{it}$

The parameters a_0, a_1, b_0, and b_1 are assumed constant over time and across firms and the disturbances v_{it} and η_{it} are each assumed to be normally distributed with mean zero and constant variance. Substituting Model 3 and Model 4 into Model 1 gives:

Model 5 $Y_{it} = (a_0 + a_1 W_{it}) + (b_0 + b_1 Z_{it})X_{it}$
$\qquad\qquad + v_{it} + \eta_{it}X_{it} + u_{it}$
$\qquad\quad = a_0 + \gamma_{it}X_{it} + \epsilon_{it}$

where $\gamma_{it} = (b_0 + b_1 Z)$ and $\epsilon_{it} = \eta_{it}X_{it} + a_1 W_{it} + u_{it} + v_{it}$ is the disturbance term of the model. This formulation assumes that W_{it} and Z_{it} are not observed. Ignoring for the moment the properties of ϵ_{it}, it is evident from Model 5 that cross-sectional estimation of the relationship between Y and X will yield an estimate of a_0 and some 'average' estimate for γ_{it} (equivalently β_{it}), but the true model parameters a_1, b_0, and b_1 cannot be identified.

In summary, not only would a cross-sectional analysis fail to provide estimates of the true model parameters but inferences regarding the significance of the explanatory variables would also be biased because the disturbance variance is heteroscedastic, since X_{it} and W_{it} vary across firms.[5] This discussion therefore makes explicit Bergh's (1995) concern that a failure to account for time-varying effects limits the scope for understanding the true nature of the relationship being investigated[6] and it also underscores

Rumelt's (1991) concern that a reliance on a single year of data for testing and making inferences about a theoretical relationship is highly problematic; not only can systematic year-to-year variations dramatically change the results depending on the time period studied but, as shown above, relying on data for only a single year limits one's ability to uncover the true relationship.

Examination of Model 5 suggests possible solutions to the preceding problems. For example, suppose both Z_{it} and W_{it} are observed. Estimates of the true model parameters can then be obtained by estimating the following model in cross-section:

Model 6 $Y_{it} = a_0 + a_1 W_{it} + b_0 X_{it} + b_1 (Z_{it}X_{it})$
$\qquad\qquad + \epsilon_{it}$

Since $\epsilon_{it} = \eta_{it}X_{it} + v_{it} + u_{it}$, the disturbance variance is again heteroscedastic. Although including W_{it} and Z_{it} (the latter enters Model 6 as an interaction variable) would permit estimates of the true model parameters to be obtained, this requires the researcher to recognize that a coefficient varies either across firms or over time and to then specify the variable(s) determining such variation. The latter is problematic since the researcher is unlikely to be able to articulate the complete list of variables involved.

The above analysis indicates that when model parameters vary across firms or over time a cross-sectional analysis will fail to identify the true model parameters and will lead to parameter estimates that are not efficient (due to heteroscedasticity). Whereas the issues arising from possible time-varying effects have been noted in the literature, the possibility that model parameters vary across firms has received scant attention. But the latter is in fact highly likely for the models commonly investigated in strategic management. For example, in the context of the diversification–performance relationship, assuming parameter constancy across firms means that this relationship is assumed to be the same across firms. But this assumption is untenable if, as the strategy literature argues, a firm's performance depends fundamentally on its ability to have a

---

[5] If the relationships for α_{it} and β_{it} where deterministic, i.e. $v_{it} = 0$ and $\eta_{it} = 0$ in Models 3 and 4, then heteroscedasticity would still arise due to the presence of W_{it} in the disturbance. In this case the problem is identical to the case of an omitted explanatory variable. We also remark that heteroscedasticity would remain an issue even if α and β, and hence W and Z, vary only over time or only across firms. However, as noted, heteroscedasticity would not arise if the relationships for α_{it} and β_{it} where deterministic.

[6] Bergh (1995) states that strategy researchers 'have not tested how yearly shifts and changes between diversification and performance are related,' and as a result, they 'have left the

longitudinal relationship between diversification and performance largely untested'.

distinctive, sustainable competitive advantage which derives from the possession and utilization of unique, non-imitable, non-transferable, firm-specific resources (Barney, 1991; Peteraf, 1993; Wernerfelt, 1984). Hence, as Barney (1986) notes, identical diversification strategies would not be expected to result in identical performance outcomes across firms.[7] In terms of Model 2, these remarks imply that β, and possibly α, differ across firms.

Extrapolating from the above argument, many of the relationships commonly examined in strategy research, such as those relating strategic choices to organizational outcomes or a firm's strategic response to changes in its competitive environment, can be expected to have parameters that are firm specific. Detection of the true parameters that characterize such relationships then requires either a time-series analysis with respect to each firm or a pooled time-series, cross-sectional analysis that allows for firm-specific differences in model parameters.

In summary, the limited ability of a cross-sectional regression analysis to account for coefficient variation across firms is an important limitation of this methodology in the context of strategic management research. While interaction variables could be used to allow for firm-specific variation in model parameters, the survey of recent empirical research indicates few attempts to use this method. In addition, parameter variation across firms or over time engenders heteroscedasticity. However, few of the recent studies in the empirical strategy literature even test for the presence of heteroscedasticity. What is needed is greater recognition on the part of researchers that these statistical issues exist and can lead to severe biases. More appropriately, strategy researchers need to adopt alternative methodologies that can alleviate the potential biases introduced if model coefficients vary across firms and over time.

MODELING PARAMETER VARIATION: AN ILLUSTRATION OF POOLED DATA ANALYSIS

Many of the problems that can arise when model parameters vary across firms or over time can be addressed by the use of pooled time-series, cross-sectional data and associated statistical methods.[8] This section offers an illustration of this methodological course of action in a specific research context.

The theoretical model investigated is the linkage between firm diversification strategy and international competition as measured by industry import penetration. Foreign competitors represent the addition of a new set of players in the industry resulting in a heightened level of competition which left unheeded will result in a deterioration of competitive position (Porter, 1980). It is hypothesized that increasing import penetration makes an industry more competitive and less attractive to the firm. Researchers have posited that firms respond to weak performance or strategic position by moving out of current businesses (Hopkins, 1991; Ravenscraft and Scherer, 1987). We therefore expect that international competition in the industries in which a firm competes will result in a defensive response on the part of the firm. The firm's competitive position across its portfolio of businesses is unlikely to be sustainable in light of significant foreign-based competition. As a result, the firm will choose to exit or decrease its investment in those businesses where it cannot maintain its competitive position. This will result in a greater strategic focus and reduced diversification; that is, we expect the level of firm diversification will be negatively related to the level of international competition in the industries in which the firm competes. This hypothesis is investigated in the context of the following model of firm diversification:[9]

---

[7] Resource-based theorists make implicit that every firm is unique due to its combination of distinctive competencies acquired and developed over time. In reality, this distinctiveness may not actually exist and furthermore the presence of these firm-specific differences may not invalidate the existence of certain general associations between concepts across all firms.

[8] For a detailed discussion on pooling time-series, cross-sectional data see, for example, Greene (1993).
[9] While not inconsistent with the existing literature, the model used here was chosen for convenience in order to provide an example of the use of LSDV. The analysis of the linkage between a firm's diversification strategy and international competition allows us to focus on the significant methodological issues that arise from a cross-sectional empirical investigation. As such, our intent in this paper is not to fully develop or test a theory about international competition. Rather, the data and form of the model used here are for illustrative purposes.

Model 7   Firm Diversification = α_{it}

$\quad +\beta_{it1}$ (Industry Import Penetration)

$\quad +\beta_{it2}$ (Industry Capital–Labor Intensity)

$\quad +\beta_{it3}$ (Industry Advertising Intensity)

$\quad +\beta_{it4}$ (Firm Market Share)

$\quad +\beta_{it5}$ (Firm Size)$_{it} + \epsilon_{it}$

This model includes controls for industry capital–labor intensity, industry advertising intensity, firm market share, and firm size, all of which have been shown in the literature to be linked to firm diversification. Before proceeding, it should be emphasized that the above model and the analysis to be presented below are only intended to illustrate the use of pooled data techniques. A fuller treatment of the hypotheses being investigated and the appropriate form of the underlying model are the subject of ongoing research by the authors.[10]

Data and variables

Firm diversification is measured with Jacquemin and Berry's (1979) entropy measure of diversification, which is calculated as

$$\text{Firm Diversification} = \Sigma s_i \ln(1/s_i)$$

where s_i is the share of a firm's total sales in a given 4-digit SIC industry i.

The industry control variables are calculated using measures weighted by s_i to reflect a firm's relative sales in the different industries represented in its business portfolio as follows:

Industry Import Penetration = $\Sigma_i s_i \text{IP}_i$

Industry Capital–Labor Intensity = $\Sigma_i s_i \text{CL}_i$

Industry Advertising Intensity = $\Sigma_i s_i \text{ADV}_i$

where, for each 4-digit SIC industry 'i', IP_i is

the ratio of imports to total domestic purchases, CL_i is the ratio of real capital stock to total industry employment, and ADV_i is the ratio of advertising expense to revenues.

Firm market share is measured as the relative market share of the firm's businesses in each of its industries. Firm size is measured as the natural logarithm of a firm's total sales as reported by COMPUSTAT.

A sample of 200 firms was randomly selected from the 1000 largest U.S. manufacturing firms in 1980 as listed by *Fortune.* The dropping of single-business firms and data limitations resulted in a consistent sample of 133 firms in each of three years 1977, 1980, and 1983. These were the only years available in the TRINET and FTC LOB data bases. Table 1 presents means, standard deviations, and correlations for all variables.

Cross-sectional and fully pooled estimates

Table 2 shows both single-year (1977, 1980, 1983) estimates as well as the fully pooled estimates. Since the cross-sectional estimates relate to specific sample years, comparison over time provides an informal analysis of the hypothesis of coefficient stability over time (a formal test is conducted below). In this regard, the coefficients appear stable in that there are no reversals in sign. The estimate for Industry Import Penetration rises and then falls (but each estimate lies within the range of expected variation as measured by the coefficient's standard deviation) while the estimates for Firm Market Share and Firm Size rise over time. The estimate for Industry Advertising Intensity in 1977 is statistically significant only at the 12 percent level and becomes increasingly insignificant in later sample years. Finally, in each year, the overall equation is significant and the model explains between 37 and 40 percent (adjusted R^2) of the variance in diversification across firms.

The results for the fully pooled model (last column of Table 2) show that, as expected, pooling yields higher *t*-statistics due to increased degrees of freedom. But despite the increase in precision, the estimate for Industry Advertising Intensity remains statistically insignificant.

The cross-sectional and fully pooled estimates appear to imply similar inferences about the relationship between diversification strategy and

---

[10] In particular, one might conjecture that the strategic response to increased import penetration would involve lagged or other dynamic effects. In response to reviewer comments, we investigated an alternative change model. The results obtained were essentially unchanged from those presented here. Importantly, the coefficient on import penetration obtained in these alternative models remained negative and statistically significant.

Table 1. Descriptive statistics and correlation matrix: Pooled sample

	Mean	S.D.	1	2	3	4	5	6
1. Firm Diversification	1.59	0.76	1.0					
2. Industry Import Penetration	0.10	0.08	−0.21***	1.0				
3. Industry Capital–Labor Intensity	41.32	45.01	0.00	−0.02	1.0			
4. Industry Advertising Intensity	0.01	0.02	−0.09*	0.02	−0.14**	1.0		
5. Firm Market Share	0.10	0.11	−0.12*	0.15*	−0.09	0.16**	1.0	
6. Firm Size	6.10	1.23	0.46***	0.07	0.22***	0.02	0.43***	1.0

$n = 399$; *$p < 0.10$; **$p < 0.05$; ***$p < 0.001$

Table 2. Regression results for firm diversification model for 1977, 1980, 1983, and pooled sample (all model coefficients restricted to be constant over time and across firms)

	1977	1980	1983	Pooled
Industry Import Penetration	−0.20***	−0.21**	−0.20**	−0.21***
Industry Capital–Labor Intensity	−0.16*	−0.21**	−0.21**	−0.19***
Industry Advertising Intensity	−0.11	−0.07	−0.01	−0.06
Firm Market Share	−0.34***	−0.41***	−0.44***	−0.39***
Firm Size	0.65***	0.69***	0.71***	0.68***
Constant	0.0	0.0	0.0	0.0
Adjusted R^2	0.37	0.40	0.38	0.40
F-statistic	16.78***	18.56***	17.37***	53.78***

$n = 133$ for years 1977, 1980, and 1983; $n = 399$ for pooled sample. Note: Standardized beta-weights reported.
*$p < 0.10$; **$p < 0.05$; ***$p < 0.001$

international competition.[11] However, both these approaches simply impose the restriction of parameter constancy across firms and over time. A proper analysis should first test if these restrictions are statistically valid. Before turning to this issue, we note that tests for heteroscedasticity were conducted using the cross-sectional results; the specific procedure used was White's generalized test.[12] In each year, the null hypothesis of homoscedasticity could not be rejected at any reasonable level of significance. Hence, subsequent analysis maintains the assumption of homoscedastic disturbances.

Testing for parameter constancy across firms and over time

To test the assumption of parameter constancy across firms or over time it is first necessary to adopt a method for modeling such variation. The method adopted here is Least Squares Dummy

---

[11] Note that a simple average (over time) of the cross-sectional estimates for a coefficient is close to the estimate derived from the fully pooled model. Any difference in these 'average' estimates reflects mostly differences in the variance of the associated variable between sample years. Hence, similarity between the simple average and fully pooled estimate for each coefficient reflects the relative constancy of the sample variances for each variable over time (see Table 1).

[12] See Greene (1993: 392).

Variables (LSDV), which uses 'dummy' variables to model coefficient variation.[13] In this framework, testing whether variation in a parameter (or set of parameters) is statistically valid is made using the common joint F-test for parameter homogeneity (e.g., testing if $\alpha_{it} = \alpha_{it} = \ldots = \alpha_{Nt} = \alpha_t$ in Model 7).

Modeling coefficient variation using the LSDV method is very general in that the dummy variables capture model and coefficient variation across firms or over time regardless of the underlying source of the variation.[14] Hence, unlike using interaction terms, the exact variables that may determine such variation need not be specified. This generality alleviates problems associated with omitted variables bias and possible heteroscedasticity but it fails to provide information about the variables that are the underlying source of coefficient variation. Instead, the dummy variables capture the combined influence of all such variables.[15]

While LSDV permits one to model coefficient variation, the problem remains that not all the coefficients in a model can be allowed to vary both across firms and over time. As such, the dimension across which the hypothesis of parameter constancy is investigated is somewhat problematic. For illustration, we first assume that any variation across firms is entirely reflected by differences in the overall model, that is, by differences across firms in the model's intercept.[16]

Given this, it is then possible to test whether the slope coefficients vary over time. This specification leads to the following form of Model 7:

Model 8   Firm Diversification

$$= \sum_t \alpha_t D_t + \sum_i \alpha_i D_i$$

$$+ \sum_t \beta_{t1} (D_t \times \text{Industry Import Penetration})$$

$$+ \sum_t \beta_{t2} (D_t \times \text{Industry Capital–}$$

$$\text{Labor Intensity})$$

$$+ \sum_t \beta_{t3} (D_t \times \text{Industry Advertising Intensity})$$

$$+ \sum_t \beta_{t4} (D_t \times \text{Firm Market Share})$$

$$+ \sum_t \beta_{t5} (D_t \times \text{Firm Size}) + \mu_{it}$$

In Model 8, the firm-specific dummy variables (D_i) take the value 1 when the observation is for firm i and zero otherwise. Similarly, the time specific dummy variables (D_t) take the value 1 when the observation is for time period t and zero otherwise. The α_t and α_i therefore capture, respectively, variation in the overall model over time and across firms while variation over time in the slope coefficients is indicated by their t subscript.[17]

Table 3A shows the results of estimating Model 8 using the pooled sample of 399 observations.[18] Comparing these results to the cross-sectional estimates (Table 2) indicates that the latter generally show greater variability over time. This suggests that the cross-sectional estimates may be subject to an omitted variables bias. Such bias is mitigated in Model 8 since the combined

---

[13] Chapter 14 of Greene (1993) contains a comprehensive discussion of LSDV and associated procedures for testing coefficient homogeneity.

[14] LSDV is easily implemented using statistical packages such as SPSS, SYSTAT, or SAS. For example, SYSTAT includes a CATEGORY command for use with its regression procedures that automatically creates dummy variables based on different values of category variables.

[15] It may be argued that the identification of the underlying variables is less important than mitigating the biases introduced by not controlling for coefficient variation across firms.

[16] Formally, the hypotheses to be investigated will help guide the choice of assumed parameter variation. For example, a more direct analysis of the current hypothesis would forgo the analysis of time variation in the slope coefficients and instead just assume that the model's intercept varies over time and across firms and that the coefficient on Industry Import Penetration varies across firms. This model nests the hypotheses of time and firm variation in one complete model rather than proceeding sequentially, as done here, to examine first time and then firm variation. In fact, we also conducted the above 'direct' analysis. As found here, significant firm variation was evident. In particular, the coefficient linking import penetration to firm diversification varies significantly across firms.

[17] For each explanatory variable we constructed three new variables—the value taken by each is the value of the explanatory variable corresponding to each year.

[18] Although flexible, LSDV uses up considerable degrees of freedom owing to the potentially large number of parameters to be estimated. An alternative method that reduces this degrees of freedom penalty is the random effects or error components model. This model interprets firm- and time-specific intercepts as random variables rather than 'fixed effects' so that these terms become part of the disturbance term. See Greene (1993) for more discussion of the random effects model.

Table 3A. Regression results for unrestricted firm diversification model using pooled data sample (coefficients allowed to vary over time or across firms)

Intercepts
Overall   0.0***
1980   0.10
1983   0.14

Industry Import Penetration
1977   −0.16***
1980   −0.17***
1983   −0.15***

Industry Capital–Labor Intensity
1977   −0.52***
1980   −0.53***
1983   −0.52***

Industry Advertising Intensity
1977   −0.26***
1980   −0.26***
1983   −0.22***

Firm Market Share
1977   −0.25***
1980   −0.24***
1983   −0.22***

Firm Size
1977   0.54***
1980   0.52**
1983   0.50**

Adjusted R^2 0.91
F-statistic   27.92***

Note: Standardized beta-weights reported. $n = 399$; *$p < 0.10$; **$p < 0.05$; ***$p < 0.001$

influence of omitted variables that vary across firms is captured by the firm-specific intercepts (α_i) and hence removed from the regression disturbance.

The statistical importance of allowing variation in the overall model and in the individual coefficients is evaluated in Table 3B.[19] The F-statistic values shown in the first column are for testing the null hypothesis that the indicated coefficients are jointly constant either over time or across firms. The *F*-values in the second column are for testing the null hypothesis that the indicated

coefficients are jointly equal to zero. Rejecting the latter hypothesis means that the variables associated with these coefficients are jointly significant for explaining a portion of the variance in Firm Diversification. From the results in column 1, the null hypothesis of coefficients stability over time can only be rejected for Industry Capital–Labor Intensity. Similarly, we cannot reject the hypothesis that the overall model is stable across firms. From the results in column 2, the null hypothesis that each set of coefficients is jointly equal to zero can be rejected in all cases except for the time-specific intercepts.

Testing for parameter variation across firms in one explanatory variable

The statistical tests in Table 3B indicate that only the coefficient on Industry Capital–Labor ratio can be taken to be invariant over time whereas the overall model varies across firms.[20] Imposing these restrictions (i.e., that all coefficients except that for the Industry Capital–Labor ratio are stable over time), we can then explore if the coefficient on Import Penetration varies across firms. This is an important issue. If this coefficient does vary across firms it would imply that a firm's Diversification strategy in response to changes in Industry Import Penetration depends on characteristics specific to that firm. This finding would then preclude making generalizations about how 'firms' would respond to increased Industry Import Penetration.

Using the LSDV approach, the hypothesis that the relationship between Firm Diversification and Industry Import Penetration is firm specific is modeled by replacing the time-specific dummy variables for Industry Import Penetration in Model 8 with firm-specific dummy variables. The model to estimate then takes the form:

Model 9   Firm Diversification $= \alpha + \sum_i \alpha_i D_i$

$$+ \sum_i D_i \beta_t \text{ (Industry Import Penetration)}$$

---

[19] All statistical packages include the option to test the equality of different coefficients. For example, in SYSTAT, the EFFECT command is used for this purpose.

[20] If additional time-series observations were available it would be possible to increase the efficiency of estimation with respect to Model 8 by exploiting possible nonzero contemporaneous covariances.

Table 3B. Tests of parameter constancy and joint significance: Unrestricted firm diversification model using pooled data sample

	Parameter constancy Value of F-statistic	Joint significance Value of F-statistic
Constancy over time		
Explanatory variables ($\beta_t = \beta$ for all t)		
Industry Import Penetration	0.26	5.66***
Industry Capital–Labor Intensity	3.32**	10.35***
Industry Advertising Intensity	0.14	6.89***
Firm Market Share	0.02	17.81***
Firm Size	0.31	2.55*
Time-specific intercepts ($\alpha_t = \alpha$ for all t)	0.15	1.07
Constancy across firms		
Firm-specific intercepts ($\alpha_i = \alpha$ for all i)	17.82***	17.76***

$n = 399$; *$p < 0.10$; **$p < 0.05$; ***$p < 0.001$

$$+ \sum_t D_t \beta_{t2} \, (\text{Industry Capital–}$$

$$\text{Labor Intensity})$$

$$+ \beta_{t3} \, (\text{Industry Advertising Intensity})$$

$$+ \beta_{t4} \, (\text{Firm Market Share})$$

$$+ \beta_{t5} \, (\text{Firm Size}) + u_{it}$$

Table 4A lists the results of estimating Model

Table 4A. Regression results for unrestricted firm diversification model using pooled data sample (coefficient on import penetration allowed to vary across firms)

Intercept	0.0*
Industry Capital–Labor Intensity	
1977	–0.46***
1980	–0.45***
1983	–0.44***
Industry Advertising Intensity	–0.79***
Firm Market Share	–0.35***
Firm Size	0.23**
Adjusted R^2	0.93
F-statistic	19.6***

Note: Standardized beta-weights reported. $n = 399$;
* $p < 0.10$; **$p < 0.05$; ***$p < 0.001$

9 using the pooled sample of 399 observations. Comparing these results to those in Table 3A, we note that allowing the coefficient on Industry Import Penetration to vary across firms leads to an increase in the relative importance of Industry Advertising Intensity and in Firm Market Share and a reduction in the importance of Firm Size.

Table 4B lists the F-statistic values for testing parameter constancy either across firms or over time. Again, we can reject the hypothesis that the coefficient on Industry Capital–Labor ratio is constant over time. The null hypothesis that the coefficient for Industry Import Penetration is constant across firms can be rejected, as can the null hypothesis that the overall model is independent of firm-specific effects.[21] From the results in column 2, the hypothesis that each set of coefficients is jointly equal to zero can be rejected in all cases.

In summary, the findings of this analysis indicate statistically significant variation across firms in the relationship linking firm diversification strategy and international competition. To the extent the model and data presented here are representative of those commonly used in stra-

---

[21] While the coefficients for Industry Import Penetration are specific to each firm we can nonetheless summarize the average effect across firms. In this regard, for those estimates significant at the 10 percent level, the average value (across firms) of the coefficient for Industry Import Penetration was –0.12; the average for all estimated coefficients was –0.06. In addition, approximately 83 percent of the significant estimates were negative. Hence, for the 'average' firm, Diversification is negatively related to Industry Import Penetration.

Table 4B. Tests of parameter constancy and joint significance: Unrestricted firm diversification model using pooled data sample

	Parameter constancy Value of F-statistic	Joint significance Value of F-statistic
Constancy over time		
Explanatory Variables ($\beta_t = \beta$ for all t)		
Industry Capital–Labor Intensity	4.91***	6.09***
Constancy across firms		
Coefficient for Industry Import Penetration	1.42**	1.55***
Firm-Specific Intercepts ($\alpha_i = \alpha$ for all i)	3.01***	3.02***

$n = 399$; *$p<0.10$; **$p<0.05$; ***$p<0.001$

tegic management research, the results suggest caution is warranted in accepting the findings of studies based on a single-year, cross-sectional analysis regarding the response of firms to their competitive environment. Instead, one should attempt to assess the possibility of variation either across firms or over time through formal statistical procedures as demonstrated here.

For the Diversification model studied here, a key finding is that the overall regression and the coefficient for Industry Import Penetration differ significantly across firms. If this finding were to hold in a more detailed analysis then it would mean that no general statement is possible about the Diversification response of 'a firm' to changes in Industry Import Penetration. Further research would then be needed to uncover the characteristics responsible for the apparent differences across firms regarding the choice of Diversification strategy in response to changes in Industry Import Penetration.

CONCLUSION

The structure–strategy–performance paradigm that is the focus of much empirical research in strategic management has both firm-specific and time-related characteristics. Yet, as illustrated by a focused survey of recent strategy research, the predominant form of analysis used to examine this paradigm is based on cross-sectional methods that assume model parameters are constant across firms and over time. Hence, the empirical methods commonly used in strategy research are glaringly at odds with the theoretical paradigm that is the focus of this research. The survey

also indicated that few studies have adequately addressed the issues of coefficient variation and heteroscedasticity. These inadequacies raise serious questions about the validity of the inferences offered by these studies.

In response to these issues, this paper examined the analytical and statistical issues that arise in a cross-sectional regression framework when model parameters vary across firms or over time. This analysis reinforced the concerns recently expressed in the literature (Bergh, 1995; Hill and Hansen, 1991; Lubatkin and Chatterjee, 1991; Rumelt, 1991) that most empirical research in strategic management has, by its reliance on cross-sectional methods, failed to adequately address the possibility that the relationships investigated might vary systematically over time. But the paper also argued that an equally if not more important concern is a failure to account for variation in model parameters across firms. The importance of this issue is not only the potential for biased inferences when cross-sectional methods are used, but, if these relationships do have firm-specific components, then no statement about the general applicability of these relationships across firms is possible. Instead, the hypothesized linkages between firm strategies and firm outcomes, or competitive environment and strategic choices, depend importantly on characteristics specific to each firm.

Given the potential for biased inferences when cross-sectional methods are used to study the relationships commonly investigated in strategy research, the paper then argued that many of the limitations of a cross-sectional analysis can be overcome by adopting analytical methods that use pooled time-series and cross-sectional data. The

application of one such method was then illustrated in the context of a specific research question: the firm's choice of diversification strategy in response to increased international competition. The results of this analysis indicated the importance of firm-specific components for the relationship examined.

Overall, the analysis of this paper points to specific recommendations that would benefit future empirical research in strategic management. First, data samples need to be broadened to include a time, as well as a firm dimension. Incorporating additional years of data on firms would seem an easy task since most of the studies surveyed relied on secondary sources of data. Subsequent application of pooled data methods then permits tests of parameter constancy across firms or over time as well as allowing one to account for the possible omission of important explanatory variables.

Empirical studies also need to give greater attention to the statistical issues that arise when the assumptions of the standard linear regression model are violated. Of particular concern is heteroscedasticity; less than 11 percent of the studies surveyed even considered this issue, despite the fact that most of these studies relied exclusively on cross-sectional data.

In raising these methodological issues it is hoped that empirical research in strategic management can more consistently address the limitations of current statistical methods and strengthen the validity of future empirical investigations.

ACKNOWLEDGEMENT

REFERENCES

Barney, J. B. (1986). 'Strategic factor markets: Expectations, luck and business strategy', *Management Science*, **32**, pp. 1230–1241.

Barney, J. B. (1991). 'Firm resources and sustained competitive advantage,' *Journal of Management*, **17**, pp. 99–120.

Bergh, D. D. (1995). 'Problems with repeated measures analysis: Demonstration with a study of the diversification and performance relationship,' *Academy of Management Journal*, **38**(6), pp. 1692–1708.

Bergh, D. D. and G. F. Holbein (1997). 'Assessment and redirection of longitudinal analysis: Demonstration with a study of the diversification and divestiture relationship,' *Strategic Management Journal*, **18**(7), pp. 557–572.

Chakravarthy, B. (1986). 'Measuring strategic performance,' *Strategic Management Journal*, **7**(5), pp. 437–458.

Godfrey, P. and C. Hill (1995). 'The problem of unobservables in strategic management research,' *Strategic Management Journal*, **16**(7), pp. 519–534.

Greene, W. H. (1993). *Econometric Analysis*, (2nd ed). Macmillan, New York.

Hill, C. W. L. and G. S. Hansen (1991). 'A longitudinal study of the cause and consequences of changes in diversification in the U.S. pharmaceutical industry 1977–1986,' *Strategic Management Journal*, **12**(3), pp. 187–199.

Hopkins, H. D. (1991). 'Acquisition and divestiture as a response to competitive position and market structure,' *Journal of Management Studies*, **28**, pp. 665–676.

Jacquemin, A. P. and C. H. Berry (1979). 'Entropy measure of diversification and corporate growth,' *Journal of Industrial Economics*, **27**, pp. 359–369.

James, W. and C. Stein (1961). 'Estimation with quadratic loss.' In J. Neyman (ed.), *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability,* Vol. 1. University Press, Berkeley, CA.

Ketchen, D. and C. Shook (1996). 'The application of cluster analysis in strategic management research: An analysis and critique,' *Strategic Management Journal*, **17**(6), pp. 441–458.

Lubatkin, M. H. and S. Chatterjee (1991). 'The strategy-shareholder value relationship: Testing temporal stability across market cycles,' *Strategic Management Journal*, **12**(4), pp. 251–270.

Pant, P. N. and W. H. Starbuck (1990). 'Innocents in the forest: Forecasting and research methods,' *Journal of Management*, **16**(2), pp. 433–460.

Peteraf, M. (1993). 'The cornerstones of competitive advantage: A resource-based view,' *Strategic Management Journal*, **14**(3), pp. 179–191.

Porter, M. E. (1980). *Competitive Strategy*. Free Press, New York.

Ravenscraft, D. and F. M. Scherer (1987). *Mergers, Selloffs, and Economic Efficiency*. Brookings, Washington, DC.

Robins, J. and M. F. Wiersema (1995). 'A resource-based approach to the multibusiness firm: Empirical analysis of portfolio interrelationships and corporate financial performance,' *Strategic Management Journal*, **16**(4), pp. 277–299.

Rumelt, R. P. (1991). 'How much does industry matter?,' *Strategic Management Journal*, **12**(3), pp. 167–185.

Venkatraman, N. and J. Grant (1986). 'Construct measurement in organizational strategy research: A critique and proposal,' *Academy of Management Review*, **11**(1), pp. 71–87.

Wernerfelt, B. (1984). 'A resource-based view of the firm,' *Strategic Management Journal*, **5**(2), pp. 171–180.