Article

Divided We Fall: How Ratios Undermine Research in Strategic Management

S. Trevis Certo[1], John R. Busenbark[2], Matias Kalm[1], and Jeffery A. LePine[1]

Abstract

Despite scholars' admonitions regarding the use of ratios in statistical analyses, the practice is common in management research. This is particularly true in the area of strategic management, where important variables of interest are operationalized as ratios. In this study, we employ simulations to demonstrate the implications of using ratios in statistical analyses. Our simulations illustrate that ratio variables produce inaccurate parameter estimates and can result in lower levels of statistical power (i.e., the ability to uncover hypothesized relationships). We also find that when an independent or a dependent variable is a ratio, the relationship between the independent and dependent variable fluctuates as the dispersion of the denominator changes. These fluctuations occur even when the correlations between the unscaled variables remain exactly the same. We also find that including ratios in models as control variables influences estimates of relationships between focal independent and dependent variables. This is true even when neither the independent or dependent variable is a ratio. We provide several recommendations for researchers who may be interested in avoiding the pitfalls of ratio variables.

[1]Department of Management & Entrepreneurship, W. P. Carey School of Business, Arizona State University, Tempe, AZ, USA
[2]Department of Management, Terry College of Business, University of Georgia, Athens, GA, USA

Corresponding Author:
S. Trevis Certo, Department of Management & Entrepreneurship, W. P. Carey School of Business, Arizona State University, Tempe, AZ 85287, USA.
Email: trevis.certo@asu.edu

> Empirical researchers love ratios—statisticians loathe them.
>
> —Jasienski and Bazzaz (1999, p. 321)

Scholars often use ratios to resolve challenges associated with empirical strategy research (Wiseman, 2009). Statisticians and researchers from an array of diverse disciplines, however, have long highlighted how the use of ratios as variables can confuse understanding regarding the meaning of relationships with and among constructs (Kronmal, 1993; Packard & Boardman, 1988; Pearson, 1896). Moreover, the use of ratios may introduce mathematical complexities that can influence statistical findings in unanticipated ways (Cohen, Cohen, West, & Aiken, 2003).

Wiseman (2009) highlighted the complexities that accompany the use of ratios in strategy research. He reported that 74% of all empirical articles published in *Strategic Management Journal* in 2007 included ratios. He also performed an extensive qualitative review of the problems associated with ratios and concluded that these problems "call into question 30 years of strategic management research that is largely based on the use of ratios to capture key constructs" (p. 76). Despite Wiseman's (2009) admonitions, we found that 79% of all empirical articles published in *SMJ* in 2015 included ratios in the statistical analyses. Importantly, the use of ratios is not limited to research published in *SMJ*. Indeed, more than two thirds of the articles (excluding meta-analyses and qualitative pieces) in *Academy of Management Journal* in 2015 used at least one ratio variable. Although the majority of these articles focused on strategy issues, ratios appeared in articles in other areas of management as well.

The primary objective of this article is to vividly demonstrate how using ratios in statistical analyses can result in inconsistent empirical findings that may impede theoretical development. Although Wiseman (2009) provided a conceptual and mathematical review of the problems associated with ratios, the unabated use of ratios suggests perhaps the derivations in his study have not resonated with strategy scholars. We therefore use a series of simulations—which are based on conditions consistent with empirical research in strategy—to investigate the potential differences that occur when ratios are used as dependent, independent, or control variables. Our results illustrate wildly inconsistent parameter estimates across study conditions. Moreover, our simulations allow us to examine the implications of using ratios with respect to statistical power, which also fluctuates dramatically across study conditions. We also hope to convey that it is problematic to interpret meaning from research using ratios, in part, because findings are driven by the dispersion of the ratio's denominator. This applies to individual studies as well as the aggregated body of knowledge that results from these studies. Indeed, given differences in the dispersion of a ratio's denominator across studies, inconsistencies in findings across studies are a certainty, and no substantive explanation will resolve them.

We reviewed strategy research involving ratios to guide the design of our simulations. Pearson (1896) introduced a formula to illustrate how changes in the coefficient of variation—a variable's standard deviation divided by its mean—of numerators and/or denominators may distort reported relationships among ratios (e.g., Jasienski & Bazzaz, 1999). Researchers in other disciplines have used simulations to examine such effects while using small coefficient of variation values of approximately 0.15 (e.g., Dunlap, Dietz, & Cortina, 1997). As we detail later, we surveyed strategy research and found that most variables involved in ratios (e.g., assets, sales, etc.) have coefficients of variation exceeding 1.0, with the highest values approaching 8.0.

The results of our simulations provide three fundamental insights for strategy scholars, though we stress that our results apply to scholars in other areas of management that also use ratios in their statistical analyses. First, reported effect sizes fluctuate dramatically across coefficients of variation in regression models that include ratios as dependent (e.g., return on assets [ROA]) and/or independent (e.g., R&D intensity) variables. The effects changed by over 900% in some cases and actually reversed signs in other cases. In contrast, our baseline models that included only numerators (i.e., no ratios) remained consistent across the different conditions.

Second, the use of ratios with coefficients of variation typically seen in strategy research resulted in lower levels of statistical significance as compared to our baseline models. In other words, the use of ratios can inflate standard errors, which in turn, reduces the likelihood of finding statistically significant coefficients when in fact there are effects to observe. Importantly, our finding that the use of ratios often reduces statistical power directly contradicts Wiseman's (2009, p. 104) conclusion that by using unscaled variables researchers "are likely to find fewer significant results than you would with ratio measures." Scholars who may have been discouraged by Wiseman's suggestion that the use of unscaled variables would be less likely to produce statistically significant results may take heart in our findings, which suggest the opposite is true.

Third, our simulations reveal that the inclusion of a ratio as a predictor influences the estimates and statistical tests associated with other predictors in the model. Regression models that include predictors operationalized with ratios—either independent or control variables—are widespread in strategy research. Our results indicate that the effect size and statistical significance associated with a variable may change substantially when the model also includes another variable operationalized as a ratio, and these changes appear to depend only on the dispersion of the ratio variable's denominator. Given the ubiquity of control variables in strategy research, perhaps in no other field is this issue so important.

Although the study of ratios is quantitative in nature, there are important implications for theoretical development. Strategy researchers have employed ratios to test theoretical perspectives such as dynamic capabilities (e.g., R&D intensity), agency theory (e.g., proportion of outside directors), and the behavioral theory of the firm (e.g., organizational slack). Our results suggest, though, that results of tests of relations among many of these constructs may be driven by variable distributions rather than the hypothesized theorized mechanisms. The implications for theory are not limited to strategy. Indeed, our findings may also apply to other areas of management, such as the study of diversity, where ratios are widely used to measure key concepts (Harrison & Klein, 2007). We hope this article serves as a catalyst toward better empirical tests and, in turn, superior theoretical understanding.

Ratios are used extensively in business and business education, so it is understandable that our message may be met with skepticism. Interestingly, this type of reaction is not limited to strategy scholars. For example, Kronmal (1993, p. 391) lists these common justifications made by statisticians for using ratios despite their understanding of the problematic mathematical properties: "Everyone else uses a ratio," "The ratio may provide a better model," "Ratios are simpler," and "The ratio is the natural quantity of interest." Although we do not dispute the usefulness of ratios in describing relevant information regarding firm strategies and performance, the point of this article is to highlight unequivocal limitations of ratios in the context of *interpreting results of statistical analysis*. Ratios may be valuable in benchmarking exercises, but the pitfalls associated with the use of ratios in regression-based statistical analyses are dramatic and inescapable. Using ratios in statistical analyses because everyone else does "sounds like what one hears from a teenager to justify particularly noxious behavior" (Kronmal, 1993, p. 391).

In the end, we are hopeful that our findings will encourage management scholars to limit the use of ratio variables in statistical analyses. Indeed, one of our main findings involves the fact that the use of ratios often reduces statistical power, and this is especially true when the ratio is used as a dependent variable (e.g., ROA, ROS, market-to-book, etc.). In other words, researchers may be *more likely* to confirm hypothesized relationships that should be confirmed if they are willing to consider easy to implement alternatives to ratio variables that we discuss later.

Problems With Ratios

Despite the reliance on ratios by strategy researchers, the literature examining the potential problems associated with ratios enjoys a long history that spans multiple disciplines such as biology (Jasienski & Bazzaz, 1999), anthropology (Smith, 2005), biochemistry (Packard & Boardman, 1988),

physiology (Curran-Everett, 2013), ecology (Liermann, Steel, Rosing, & Guttorp, 2004), and statistics (Kim, 1999). The use of ratios extends to other fields of organizational studies, such as accounting (Brown, Lo, & Lys, 1999) and finance (Lev & Sunder, 1979). Wiseman (2009) provides an accessible overview of these problems. Rather than reproducing Wiseman's arguments, our intent here is to clarify and supplement them.

Spurious Relationships Owing to Ratios

As early as 1896, Pearson began to question the potential spuriousness that results from using ratios. He suggested that transforming data into ratios may result in relationships that might not otherwise exist. Simply stated, he suggested that statistical relationships among ratios might be due to relationships between numerators, relationships between the denominators, or both. Pearson's concern with ratios stimulated scholars to think about problems associated with using ratios in more complicated analyses such as regression. Perhaps the most well-known example is Neyman's (1952) illustration of the problems that may occur when using ratios as both dependent and independent variables. His (fictitious) example involved understanding how the presence of storks in a county was correlated with the number of births in a county. Using regression, he found that the number of storks per capita was a significant (and positive) predictor of the number of births per capita. However, when using the number of storks as a predictor of the number of births while controlling for county population, the effect of storks on subsequent births disappeared. In this example, the per capita relationship was driven by the fact that the independent and dependent variables shared the same denominator.

Scholars have also noted that ratios obfuscate the nature of the focal relationship. Kronmal's (1993) example using body mass index (BMI) as a predictor of waist size illustrates this concern. Kronmal characterized BMI, which measures the ratio of weight to height squared, as "probably the most widely used ratio in medical research" (p. 388). Although the regression showed that BMI is a significant predictor of waist size, Kronmal recognized the difficulty in ascertaining whether the reported effect of BMI was driven by the numerator, the denominator, or both. To answer this question, he ran another regression of waist size on weight and height squared as separate independent variables. He found that "it is principally weight that determines waist size, with the height making a lesser contribution" (Kronmal, 1993, p. 388).

Another intuitive way to understand the issues associated with ratios involves recognizing that a ratio represents the product of the numerator and the inverse of the denominator (Kronmal, 1993). Put differently, the ratio of x over (or divided by) z is created by multiplying x and $1/z$. That is, the ratio "x/z" and the product terms "$x*1/z$" are mathematically identical. For instance, ROA can be thought of as the product of net income and the inverse of assets (i.e., 1/assets). This simple mathematical restatement provides a vantage point for understanding the underlying complexity of a ratio in the context of multiple regression. As a simple example, consider that operationalizing performance as ROA as an independent variable in a regression model is precisely the same thing as specifying a model with a product term that carries an interaction without entering the lower order terms—net income and 1/assets. Numerous treatments outline requirements for including main effects in models testing interactions (Aiken & West, 1991; Edwards, Lance, & Vandenberg, 2009; Kennedy, 2008). Omitting the main effects from a model that includes interaction terms may create an omitted variable problem, which makes the effects of an interaction difficult—if not impossible—to interpret (Brambor, Clark, & Golder, 2006).

The Importance of Variable Dispersion: Coefficient of Variation

While the discussion above highlights general limitations of ratio variables, scholars have also sought to identify factors that explain how and under what conditions ratios are most problematic.

Among the first to focus on this issue was Pearson (1896), who suggested that the correlation between two ratios depends on the dispersion of the numerators and denominators. In this regard, he highlighted the role of the *coefficient of variation*, which refers to the standard deviation of a variable divided by its mean. The coefficient of variation provides information regarding a variable's distribution, with higher values representing larger degrees of dispersion. Pearson developed a formula to approximate how the coefficients of variation influence the correlations between ratios. He used this formula to illustrate that the correlation between two ratios may vary based only on the distributions of the numerator and denominator.

Dunlap et al. (1997) employed a series of simulations to examine the accuracy of this approximation. Their simulations revealed that Pearson's approximation worked well when the coefficients of variation remained small (i.e., approximately 0.15). They also found that the correlations between ratios became more misleading as the correlations between the numerators and denominators increased. Indeed, Dunlap et al. (1997) noted that "there is no way of even estimating the X-Y relationship from correlated ratios without knowing a great deal about the original variables" (p. 191). The primary conclusion of this research is that correlations between ratios depend largely on the variable distributions of the numerators and denominators. As Curran-Everett (2013) noted, "A ratio is seductively simple. There is just one problem: the analysis of a ratio is complex" (p. 213).

Ratios in Strategy Research

Although strategy scholars have used ratios to address issues such as heteroscedasticity and to study multidimensional constructs, the primary reason they use ratios is to account for firm size (Wiseman, 2009). As Wiseman (2009) pointed out in his review of ratios in strategy research, many variables in financial statements are correlated with firm size. Presumably, then, scholars use indices of firm size in the denominator of a ratio to help scale the numerator for size effects. To this end, scholars have incorporated size as a denominator to create variables such as R&D intensity, return on assets, and several other variables.

To examine how the use of ratios has changed since Wiseman's survey, we reviewed all articles published in *SMJ* in 2015. We found that of 99 empirical studies published that year, 78% included at least one ratio. In other words, strategy scholars continue to use ratios since Wiseman's work. Our review of *SMJ* articles published in 2015 also provided insight into *how* researchers used size-based ratios in their analyses. Indeed, strategy research seldom focuses on bivariate associations, and therefore, it is important to consider implications of the positioning of ratios in multiple regression models. Figure 1 provides a Venn diagram that illustrates the extent to which ratios appeared as dependent, independent, and control variables.

Of these 99 empirical studies, we found that 43% included ratios as dependent variables, 50% used ratios as independent variables, and 71% used ratios as control variables. We also found that 40% of empirical articles included ratios on both sides of their statistical models (i.e., a ratio as a dependent variable and a ratio as an independent variable and/or a control variable). In sum, our review revealed that strategy scholars continue to use ratios extensively as dependent, independent, and control variables.[1] As we demonstrate in the descriptions of our simulations, the consequences of using ratios change depending on whether ratios are used to represent dependent, independent, and/or control variables.

Variable Dispersion in Strategy Research

Despite the recognition that the dispersion of the components of a ratio may influence correlations of variables involving ratios, we have very little knowledge of how the dispersion of variables used to create ratios in strategy may influence results of statistical analyses of more complicated models that
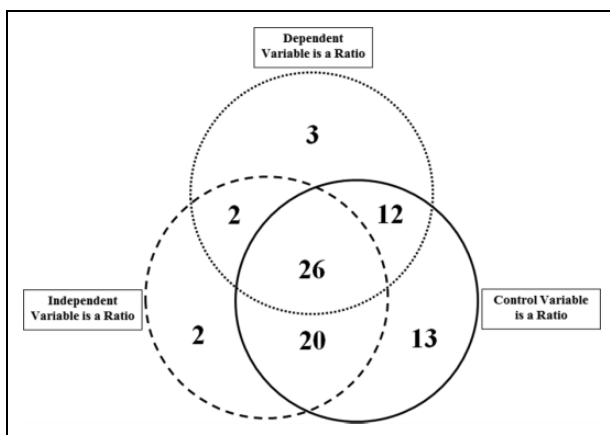
Figure 1. Venn Diagram Depicting the Use of Ratios in Different Ways. Note: Each of these numbers represents both a raw number and percentage since there were 99 empirical articles in SMJ in 2015.

management scholars typically employ. The heavy reliance of strategy research on archival data, for instance, may introduce differences in statistical results that may not apply to research on ratios in other fields (Ketchen, Boyd, & Bergh, 2008). In addition, the role of the firm as the unit of analysis may introduce differences in coefficient of variation values that do not apply to research with different units of analysis (Josefy, Kuban, Ireland, & Hitt, 2015).

To gain insight into the potential influence of variable dispersion in strategic management research, we collected data from several sources to examine variables that strategy scholars typically use as numerators or denominators in ratios. We collected data from 2015 for three samples of firms: S&P 500, S&P MidCap 400, and S&P SmallCap 600. For each variable, we calculated the coefficient of variation in each sample and also for the overall population of firms (i.e., S&P 1500). In the "Raw Data" column of Table 1, we report the coefficient of variation corresponding to each variable for the different samples. We also examine how winsorizing at the 1% level, a popular transformation technique in strategy research used to reduce the influence of outliers (e.g., Carow, Heron, & Saxton, 2004; O'Connell & O'Sullivan, 2014), influences coefficients of variation.

Two results in Table 1 warrant elaboration. First, the coefficient of variation varies dramatically across the different variables, samples, and transformations. Whereas several variables that are typically used as numerators and denominators in ratios (e.g., total assets, revenues, and R&D expenditures) have coefficients of variation that exceed 1, other variables (e.g., board size, total equity, and total CEO compensation) have coefficients of variation less than 1. Second, winsorizing influences corresponding coefficients of variation. For example, the coefficient of variation for assets in the S&P 500 sample is 4.9, while it is less than half that value (2.2) for those same firms when the data are winsorized. Although winsorizing may have benefits as it relates to outliers, its effect on the coefficient of variation may influence results associated with models incorporating ratios in ways that have yet to be considered.

Summary

Most prior research on ratio variables conducted in other disciplines focuses on relatively small coefficients of variation (typically, values less than 1) and considers the implications regarding bivariate correlations. Our review of strategic management literature reveals a much wider range of coefficients of variation. Moreover, as we alluded to earlier, analyses in strategy typically involve the use of multiple regression and related techniques where ratios are positioned in a variety of

Table 1. Coefficients of Variation and Examples of Several Popular Variables Used in Ratios.

Variable	Raw Data				Winsorized at 1%				Source		Ratio Examples
	S&P 500	MidCap	SmallCap	S&P 1500	S&P 500	MidCap	SmallCap	S&P 1500	Database	Variable	
Assets	4.9	1.4	1.8	3.2	2.2	1.3	1.3	2.7	Compustat	at	ROA, R&D intensity, and M&A size (e.g., Makri, Hitt, & Lane, 2010; Tyler & Caner, 2016)
CEO's option pay	3.2	1.9	2.0	2.6	1.4	1.6	1.8	1.8	Execucomp	option_awards_fv	Option pay ratio (e.g., Shi, Connelly, & Sanders, 2016; Wowak, Mannor, & Wowak, 2015)
CEO's total compensation	1.0	0.6	1.6	0.7	0.6	0.6	0.6	0.8	Execucomp	tdc1	Contingent pay ratio (e.g., G. Chen, 2015; Feldman, 2016)
Employees	3.4	1.7	4.5	2.4	1.5	1.6	1.8	2.1	Compustat	emp	Employee productivity (e.g., Bertrand & Capron, 2015; Chadwick, Super, & Kwon, 2015)
Long-term debt	3.5	1.3	1.7	2.2	2.0	1.0	1.4	2.2	Compustat	dltt	Debt-to-equity and financial leverage (e.g., Chizema, Liu, Lu, & Gao, 2015; Kang, 2016)
Net income	4.4	4.1	7.6	2.6	2.0	3.2	5.5	3.0	Compustat	ni	ROA, ROE, and ROIC (e.g., Krause, Priem, & Love, 2015; Vomberg, Homburg, & Bornemann, 2015)
Number of directors	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	ISS	_n	Proportion of outsiders, females, minorities, interlocking, or CEO directors (e.g., Dezső, Ross, & Uribe, 2016; Diestre, Rajagopalan, & Dutta, 2015)
Number of females on the board of directors	1.2	1.4	1.7	0.8	0.8	1.3	1.6	1.1	ISS	female	Proportion of female directors (e.g., Gupta, Briscoe, & Hambrick, 2017; Humphery-Jenner, 2014)
Number of minorities on the board of directors	0.7	0.7	0.9	0.5	0.5	0.7	0.9	0.7	ISS	ethnicity	Proportion of minority directors (e.g., Campbell, Campbell, Sirmon, Bierman, & Tuggle, 2012; Hill, Upadhyay, & Beekun, 2015)
Number of outsiders on the board of directors	0.3	0.2	0.3	0.3	0.2	0.2	0.3	0.3	ISS	classification	Proportion of outside directors (e.g., Betschinger, 2015; Zhu & Yoshikawa, 2016)
R&D-expenses	3.5	1.6	1.5	2.2	2.2	1.6	1.4	3.2	Compustat	xrd	R&D intensity (e.g., Bromiley, Rau, & Zhang, 2017; Lim, 2015)

(continued)

Table 1. (continued)

Variable	Raw Data				Winsorized at 1%				Source		Ratio Examples
	S&P 500	MidCap	SmallCap	S&P 1500	S&P 500	MidCap	SmallCap	S&P 1500	Database	Variable	
Revenue	2.8	1.3	1.3	1.8	1.5	1.2	1.2	2.2	Compustat	revt	Return on sales (ROS) and value added per sales (e.g., C.-M. Chen, Delmas, & Lieberman, 2015; Zhao & Murrell, 2016)
Shares outstanding	2.6	0.8	0.7	1.7	1.5	0.7	0.7	1.9	Compustat	csho	Earnings per share (EPS), proportion of blockholder ownership, proportion of institutional ownership (e.g., Chang & Shim, 2015; Judge et al., 2015)
Total equity	3.3	0.8	0.8	2.1	1.9	0.8	0.7	2.3	Compustat	teq	Return on equity (ROE), market-to-book, and debt-to-equity (e.g., Flammer, 2015; Shiu & Yang, 2017)

different ways. Given these differences in the ranges of dispersion and analytic approaches, it is uncertain how the insights and criticisms regarding ratios apply to strategy and the inferences we draw from our data. To address this uncertainty, we develop a series of simulations which we describe in the next section.

Simulations

General Setup

We created a series of simulations with the aim of testing many of the conditions outlined in Figure 1. Our simulations involved two general steps. In the first step, we generated data with characteristics consistent with both existing strategy research and prior simulations (e.g., Certo, Busenbark, Woo, & Semadeni, 2016; Semadeni, Withers, & Certo, 2014). In the second step, we employed different models to examine the results associated with each dataset. The first model we simulated was a baseline condition with no ratios. This model allowed us to establish the true relationships between independent and dependent variables. We then specified the simulation to examine models that include a ratio as either a dependent, independent, or control variable. Finally, we combined these previous studies to examine models including ratios as *both* dependent and independent/control variables. In the following section, we describe the data generation process, the analytical models used in the various studies, and the results of each simulation.

Data generation. We generated datasets with 1,000 (i) observations containing five variables: y (dependent variable), x_1 (independent variable 1), x_2 (independent variable 2), z_1 (scaling variable also included as a control), and z_2 (a variable used to scale the dependent variable). We set the correlations between each of the independent covariates and the dependent variable (y) at 0.15, and we set the correlations between independent covariates at 0.125. The means and standard deviations of all the variables are 1 and 2.5, respectively, but as we explain later, we altered the means of the scaling variables to create different coefficients of variations.

We chose all these parameters because they represent a realistic baseline condition in terms of effect size and variance explained in the overall model (Certo et al., 2016; Cohen, 1992; Semadeni et al., 2014) and are also consistent with organizational research on effect sizes (Bosco, Aguinis, Singh, Field, & Pierce, 2015).[2] We employed extensive sensitivity analyses on all of the parameters and values selected, and our results remain substantively similar regardless of the values we select. For instance, we varied the sample size with values ranging from 500 to 5,000, and our results remained substantively similar.

We will use an example throughout this section to highlight the intuition of this setup. Our models included a dependent variable (e.g., *net income* [y]) as a function of an independent variable that may be scaled (e.g., *R&D expenditures* [x_1]), an independent variable that remains unscaled (e.g., *M&A activity* [x_2]), and two scaling variables (e.g., *total assets* [z_1] and *total equity* [z_2]). As our simulations progress, we consider ratios calculated from these underlying variables to understand how ratios influence results of statistical analyses.

Simulation conditions. In all of our simulations, we examine how estimates vary based on the coefficient of variation of our scaling variable, z_1. To examine results at different coefficients of variation, we could alter either the means or the standard deviations of our scaling variable. Because ordinary least squares (OLS) models examine changes in variance, any changes to the standard deviations of the variables would alter overall model properties such as total variance and variance explained. As a result, we kept the standard deviations of the variables constant while adjusting the variable means. Specifically, we varied the mean of z_1 such that the coefficients of variation for this variable ranges from approximately 0.1 to 5, which is consistent with the values listed in Table 1. The coefficients of variation increase in increments of 0.1 from 0.1 through 1 and

in increments of 1 from 1 through 5 (with the exception of the value 2.5, which represents the baseline described in the setup).[3]

We should note that we performed the same technique across different coefficients of variation of our central independent variable (x_1). We did this independent of, and in addition to, changing the coefficients of variation of z_1. Doing this does not substantively change our results with respect to dramatic inconsistencies, but the outcomes become remarkably more complex. Although it is important to note that the results we describe below can derive from changes in the coefficient of variation of either the numerator or denominator, we focused solely on the denominator in this article for the sake of brevity, consistency, and ease of interpretation.

Data Analysis

Estimator. We used OLS regression in all of our studies. For each coefficient of variation condition, we ran 1,000 iterations. As we progressed throughout the studies, we examined different combinations of ratios, but in each case we were interested in the properties associated with the coefficients and how they changed with different model specifications and coefficient of variation conditions.[4]

Outcome variables. Consistent with past research on simulations in strategy (Certo et al., 2016; Semadeni et al., 2014), we relied on two primary outcomes pertaining to the parameter estimates associated with each of our variables. First, B_{Med} represents the median coefficient estimate over 1,000 simulations per condition. Second, *PerSig* represents the percentage of simulations in which the model reported a statistically significant coefficient (5% level). Deviations from the *PerSig* represent Type I and Type II errors. Type I errors represent instances when a scholar may find a statistically significant effect when in fact there is no relationship. Type II errors reflect instances when a scholar may not find a statistically significant finding when in fact there is a relationship. Put differently, Type II errors increase as statistical power decreases. Deviations from B_{Med} represent inaccuracy (or bias) in the estimates (Certo et al., 2016).

Study 1: Unscaled Dependent, Independent, and Control Variables (No Ratios)

In this baseline study, we investigate the results of OLS when the dependent, independent, and control variables are unscaled. As we mentioned, we use the example of R&D as a predictor of net income throughout these studies to provide context for how each study varies from the original general setup. In this study, we examine the example of net income (y_i) as a function of R&D (x_{1i}), M&A activity (x_{2i}) and assets as a control variable (z_{1i}).[5] This is displayed in Equation 1:

$$y_i = a + B_1 x_{1i} + B_2 x_{2i} + B_3 z_{1i} + e_i \qquad (1)$$

This approach allows us to understand how model results might vary as the distribution of the scaling variable (z_1) changes. As our baseline study, we are interested in the properties of B_1 and B_2, which estimate the effects of the independent variables on the dependent variable. It is well known that in the context of OLS regression the first derivative of a regression equation represents the marginal effect of an independent variable on a dependent variable (Kennedy, 2008; Wooldridge, 2010). In this example, then, the expected marginal effect of x_1 on y (i.e., $\frac{d_y}{d_{x_1}}$) is B_1. Consequently, we expect this baseline study to produce a parameter estimate that reflects the true relationship between the independent variables and the dependent variable.

Results. In this baseline study, we examined B_{Med} corresponding to both x_1 and x_2. OLS estimates of B_{Med} corresponding to both x_1 and x_2 remained consistent at approximately .09 as the coefficient of variation for z_1 changed. In addition, the percentage of cases in which the OLS model reported a

significant B_1 and B_2 remained consisted around approximately 80% as the coefficient of variation for z_1 changed. In other words, changes in the coefficient of variation of z_1 do not appear to alter parameter estimates or statistical power when the model has unscaled variables (i.e., no ratios).[6]

Study 2: Dependent Variable Is a Ratio

In Study 1, we found that the results of OLS remained consistent across the various coefficient of variation conditions for z_1. In the present study, we investigate the effects of using a ratio as a dependent variable. As illustrated in Figure 1, almost half of empirical research in *SMJ* in 2015 includes dependent variables operationalized as ratios. We use the same data generated in Study 1 to estimate the following equation:

$$\frac{y_i}{z_{1i}} = B_1 x_{1i} + B_2 x_{2i} + B_3 z_{1i} + e_i \tag{2}$$

In this study, we focus on the estimate for B_1 since this represents the focal independent variable of interest. To provide context, we revisit our prior example of R&D expenditures and net income. In this example, we use total assets (z_1) to scale the dependent variable, which equates to using ROA as our dependent variable.

Rewriting Equation 2 helps to illustrate how the relationship between y and x_1 may differ from our baseline study:

$$y_i = B_1 z_{1i} x_{1i} + B_2 z_{1i} x_{2i} + B_3 z_{1i}^2 + e_i z_{1i} \tag{2'}$$

The first derivative of this equation reveals that the marginal effect of x_1 on y is $B_1 z_1$ instead of B_1 in our baseline study; the relationship between x_1 and y now depends on the value of z_1.

It is also important to note that the marginal effects predicted by this equation depend on the assumptions of OLS regression being met, and there is evidence that this is not the case. First, Equation 2' includes product terms that reflect interactions and curvilinear terms while omitting the main effects of the interactions and curvilinearity (e.g., the coefficients for x_1 and z_1). Researchers in multiple disciplines have highlighted the omitted variable problems that may arise when failing to include lower order terms when examining interactions and curvilinearity (Aiken & West, 1991; Edwards et al., 2009; Kennedy, 2008).

In addition, the equation shows that the error term (ez_1) differs from that of our baseline study (e). While scholars have noted how the denominator might change the error term and lead to inconsistencies (e.g., Kronmal, 1993; Wiseman, 2009), it has not been recognized that these changes result from potential endogeneity. Specifically, z_1 is now a component of the error term and all of the independent variables, which creates this endogeneity (Kennedy, 2008; Semadeni et al., 2014). The results of our simulations illustrate the implications of these factors.

Results. Figure 2 and Figure 3 illustrate B_{Med} and *PerSig* corresponding to B_1 when the dependent variable is a ratio. In Figure 2, the estimates for B_{Med} vary drastically as the coefficient of variation of z_1 increases. Accordingly, the changes in B_{Med} across the different coefficients of variation of z_1 represent inaccuracy and bias in the model. At the same time, Figure 3 shows that *PerSig* for B_1 drops dramatically as the dispersion of z_1 increases. In fact, *PerSig* drops below 10% for coefficient of variation of z_1 greater than 0.5. In other words, the power to detect an existing relationship is less than 10% in many situations that are relevant to strategy scholars.

Study 3: Independent Variable Is a Ratio

In Study 2, we examined the effects of using a ratio as a dependent variable. In this study, we examine the effects of using a ratio as an independent variable. As shown in Figure 1, about 50% of
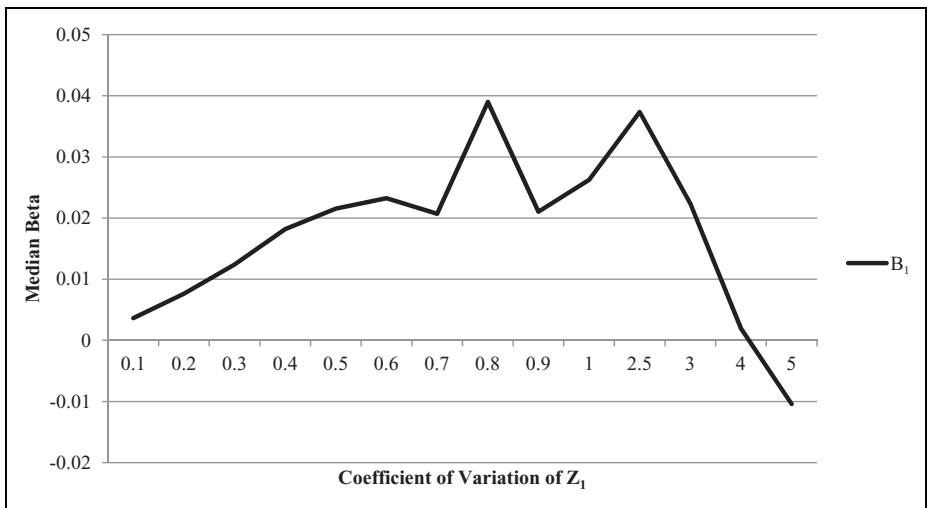
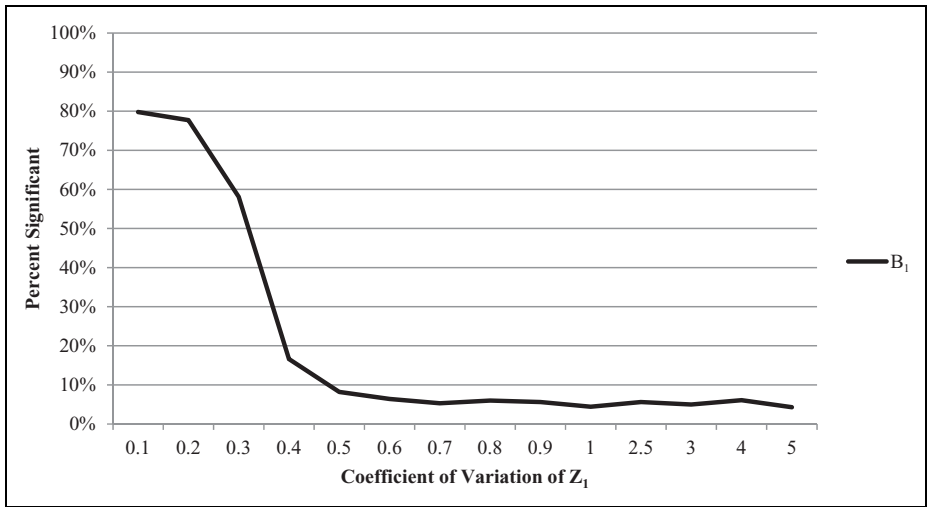**Figure 2.** Median Beta When the Dependent Variable Is a Ratio (Study 2).

Figure 3. Percentage Significant When the Dependent Variable Is a Ratio (Study 2).

empirical articles published in *SMJ* in 2015 include a ratio as an independent variable. We revisit our previous example of R&D expenditures and net income, but for this study we scale the independent variable by assets to create R&D intensity. Using the same data generated in our baseline study we estimated the following equation:

$$y_i = B_1 \frac{x_{1i}}{z_{1i}} + B_2 x_{2i} + B_3 z_{1i} + e_i \tag{3}$$

Taking the first derivative of this equation reveals that the marginal effect of x_1 on y is now $B_1\left(\frac{1}{z_1}\right)$ instead of the marginal effect of B_1 in our baseline study. Again, however, this approach may violate key assumptions of OLS regression because it includes the product term, $x_1 * \frac{1}{z_1}$. Like Study 2, this model lacks the lower order terms pertaining to this product term (i.e., both x_1 and $\frac{1}{z_1}$), which could
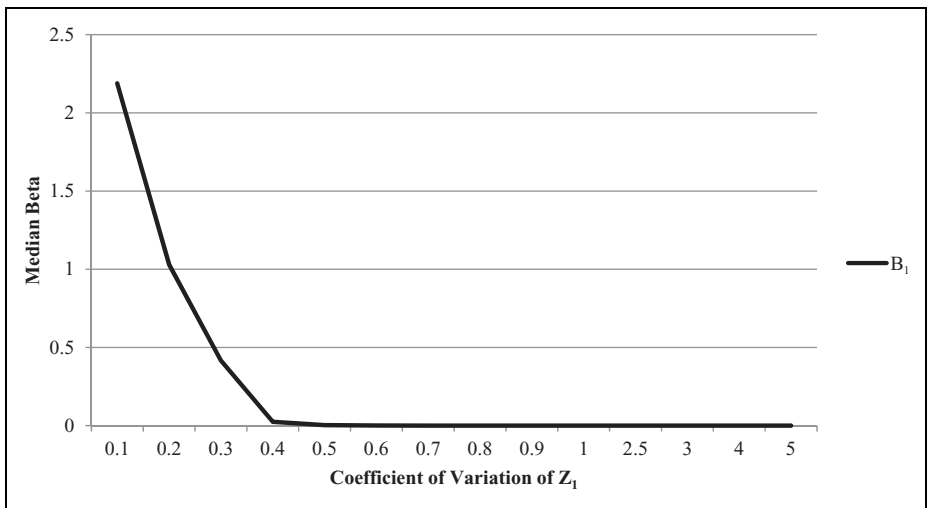
Figure 4. Median Beta When the Independent Variable Is a Ratio (Study 3).
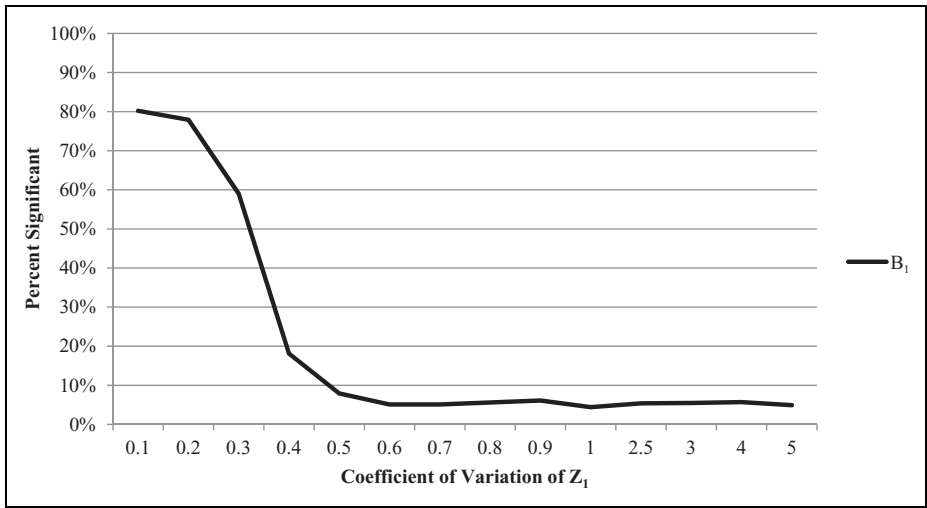

Figure 5. Percentage Significant When the Independent Variable Is a Ratio (Study 3).

create an omitted variables problem. Our simulations allow us to investigate the implications of this issue in more detail.

   Results. Figure 4 shows how the coefficient of variation influences B_{Med} for the scaled independent variable, B_1 (e.g., R&D intensity). When the coefficient of variation of z_1 is 0.1, B_{Med} is much higher than when the coefficient of variation of z_1 is 5. In fact, B_{Med} drops to near zero when the coefficient of variation of z_1 is approximately 0.4, and it remains at these low values as the coefficient of variation increases. Figure 5 illustrates a corresponding pattern as *PerSig* declines dramatically as the coefficient of variation of z_1 increases. In fact, this value falls to approximately 7% when the coefficient of variation of assets is 0.5 and remains at low levels as the coefficient of variation increases.

Study 4: Both Dependent and Independent Variables Are Ratios

Studies 2 and 3 illustrate the inconsistencies that may occur when scaling either the dependent or independent variables by another variable. With Study 4, we extend our investigation to examine the effects of using ratios as *both* independent and dependent variables. According to our review, approximately 30% of empirical studies published in *SMJ* in 2015 include ratios as *both* dependent and independent variables (these independent variables were either with or without additional ratios as controls).

Our review of strategy research indicated that the denominators of the independent and dependent variables frequently differed. Consequently, in this study we included both scaling variables, z_1 and z_2. We then used OLS to obtain estimates of the coefficients and standard errors needed to calculate our outcome variables based on the following equation:

$$\frac{y_i}{z_{2i}} = B_1 \frac{x_{1i}}{z_{1i}} + B_2 x_{2i} + B_3 z_{1i} + e_i \tag{4}$$

In this study, we are again interested in B_1. We created four different correlations between the two denominators, z_1 and z_2: 0.25, 0.50, 0.75, and 1.0. In each condition, we kept the coefficients of variation of z_1 and z_2 identical to ensure our results were driven by different correlations between the denominators and not different dispersions. Extending our prior example, this study involved R&D intensity as the primary ($\frac{x_{1i}}{z_{1i}}$) independent variable. Using both z_{1i} and z_{2i} allowed us to extend this study to dependent variables with different denominators such as return on equity (ROE), capital intensity, and so on. For this example, we refer to our second scaling variable (z_2) as total equity, which means the dependent variable reflects ROE. The two denominators are perfectly correlated in the final condition, similar to contexts in which the same variable is used for both denominators (e.g., ROA and R&D intensity).

Like our prior studies, rewriting this equation can help to better understand the marginal effects of the independent variable in this study:

$$y_i = B_1 \frac{z_{2i} x_{1i}}{z_{1i}} + B_2 z_{2i} x_{2i} + B_3 z_{1i} z_{2i} + e_i z_{2i} \tag{4'}$$

Taking the first derivative shows that the marginal effect of x_1 on y is $B_1 \frac{z_2}{z_1}$ instead of B_1, as in our baseline study. Again, however, the ability of OLS to provide unbiased parameter estimates depends on several assumptions, and there is reason to question some of these assumptions. In fact, this equation—which includes ratios as both independent and dependent variables—introduces the problems of both previous studies. Rewriting this equation shows that it now includes product terms that reflect a three-way interaction ($z_2 * x_1 * \frac{1}{z_1}$) and two two-way interactions ($z_2 * x_2$ and $z_1 * z_2$). Because this equation does not include any of the lower-order terms, omitted variable problems could result. Furthermore, Equation 4' shows z_2 is a component of each independent variable and the error term, which results in endogeneity (Semadeni et al., 2014). Our simulations allow us to investigate the implications of these issues.

Results. Figures 6 and 7 show the results when the dependent and independent variables are ratios (as in Equation 4). Because this study involves two denominators, we change our figures slightly from our previous studies. In each figure, we depict how the outcome variable changes across different correlations between the scaling variables. Figure 6 shows that B_{Med} is biased and tends to increase as the coefficient of variation of the denominators increases. *PerSig* for B_1 in Figure 7 shows a similar pattern to B_{Med} in Figure 6.
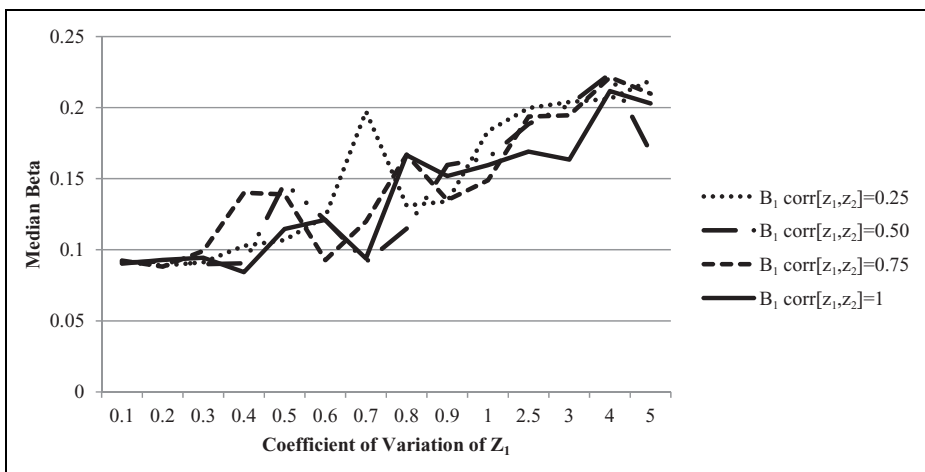
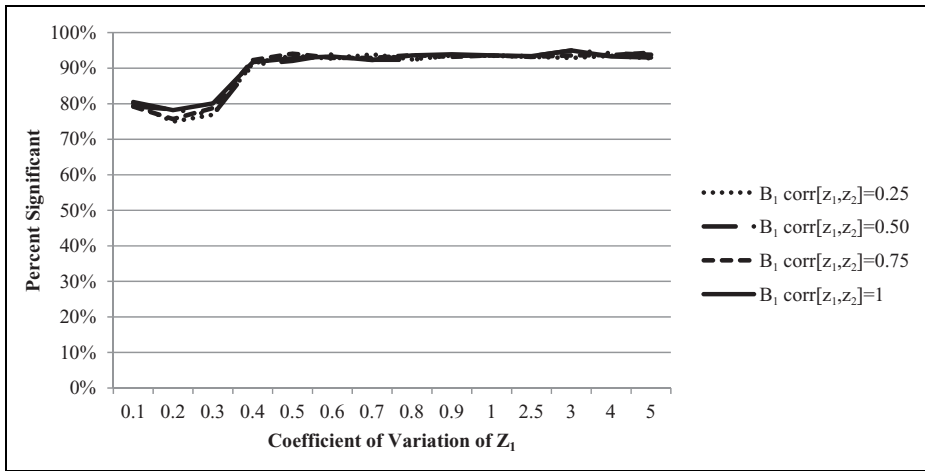**Figure 6.** Median Beta When the Independent and Dependent Variables Are Ratios (Study 4).

Figure 7. Percentage Significant When the Independent and Dependent Variables Are Ratios (Study 4).

Study 5: When a Control Variable Is a Ratio

Our previous studies examined how varying distributions influence analytical results when ratios are used as either the independent and/or dependent variables. In the present study, we extend our analysis to examine how the use of a ratio predictor might influence the estimates for other non-ratio variables in a regression. Indeed, our review of *SMJ* in 2015 reveals that 33% of articles include a ratio as a control variable when the dependent variable is not a ratio.

Returning to our example may help to illustrate the intuition of this study. Suppose that a researcher is interested in investigating the influence of M&A activity (x_2) on net income while including R&D intensity and assets as control variables. To examine this issue, we use the estimates obtained from Equation 3 in Study 3, which we reproduce for convenience:

$$y_i = B_1 \frac{x_{1i}}{z_{1i}} + B_2 x_{2i} + B_3 z_{1i} + e_i \tag{3}$$
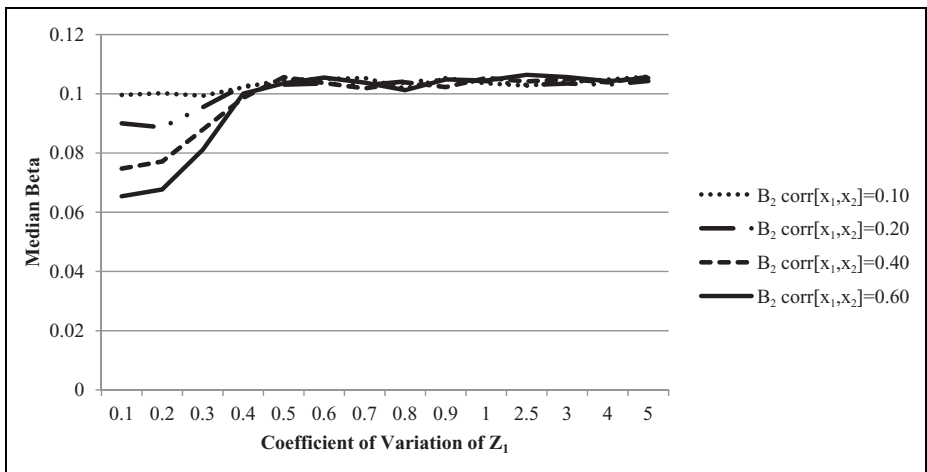
Figure 8. Median Beta When a Control Variable Is a Ratio (Study 5).

In Study 3, we found that the use of the ratio influenced B_1, but in the present study we investigate the impact on B_2. Specifically, we now focus on how estimates for B_2 fluctuate across different coefficients of variation for z_1, even though x_2 remains unscaled. We are also concerned with how the correlation between x_1 and x_2 influences the estimates for B_2. Stated differently, our concern here is whether hypothesized non-ratio variables are innocent bystanders when ratios are included in the model, such as when they are used as controls.

Taking the first derivative of this equation shows that the marginal effect of x_2 on y is simply B_2. This implies that the marginal effect should remain constant. At the same time, though, the equation lacks the main effects (i.e., x_1 and $\frac{1}{z_1}$) of the product term ($\frac{x_1}{z_1}$), and it is possible that this will influence the error term and, consequently, the parameter estimates. Our simulations will allow us to investigate the implications of these factors in more detail.

Results. As shown in Figure 8, B_2 is highly dependent on the coefficient of variation for z_1, even though z_1 is only included in a ratio involving x_1. As the coefficient of variation of z_1 increases, so too does the estimate for B_2. This relationship is accentuated by the correlation of x_1 and x_2. As the correlation between these two variables increases, B_2 also increases at a higher rate. *PerSig* for B_2 in Figure 9 shows a similar pattern to B_{Med} in Figure 8. The estimate for B_2 is statistically significant more frequently when the coefficient of variation of assets increases. In addition, *PerSig* increases at a higher rate when the correlation between x_1 and x_2 increases.

Comparing these results to those of Study 3 provides intuition underlying these relationships. Study 3's Figure 4, which used the same equation but examined a different coefficient, shows that the effect of the ratio *decreased* as the coefficient of variation for z_1 increased. This was because x_1 was omitted from the model. In this study, though, the reported effect of the non-ratio variable *increased*, and we suggest that it is because x_2 now reflects the variance of the same omitted variable.

Study 6: When Control and Dependent Variables Are Used as Ratios

In the previous study, we examined how using a ratio as a control variable might influence the effects of an independent variable on a dependent variable. In this study, we extend the analysis to understand the effects when the dependent variable is also a ratio. According to our review,
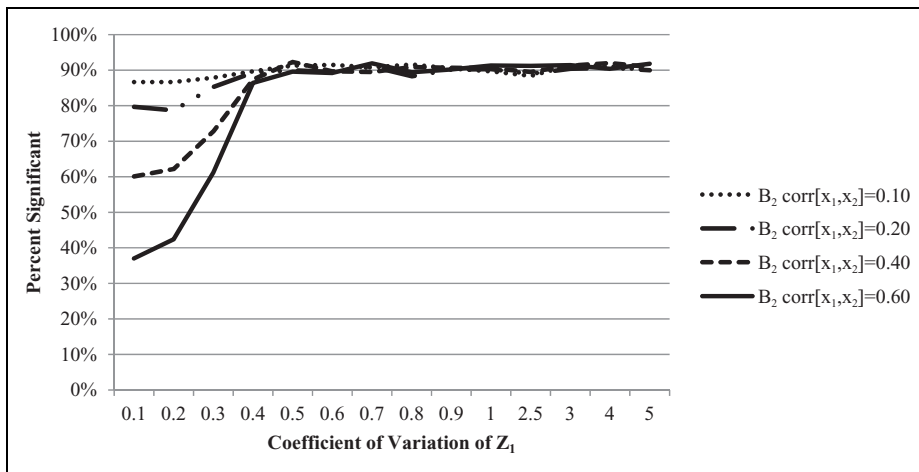
Figure 9. Percentage Significant When a Control Variable Is a Ratio (Study 5).

approximately 40% of empirical studies included ratios as both control and dependent variables. Again, extending our example may help to illustrate the intuition of this study. Suppose a researcher is interested in investigating the influence of M&A activity on ROA while controlling for a ratio such as R&D intensity. To test this, we use the estimates obtained from Equation 4 in Study 4, which we reproduce for convenience:

$$\frac{y_i}{z_{2i}} = B_1 \frac{x_{1i}}{z_{1i}} + B_2 x_{2i} + B_3 z_{1i} + e_i \qquad (4)$$

In Study 4 we were interested in B_1, but in this study we are interested in how the use of a ratio influences B_2. Like the previous study (Study 5), we examine how the correlation between x_1 and x_2 influences the results.

To understand the marginal effects of the independent variable $(\frac{x_1}{z_1})$ on the dependent variable, it is once again necessary to rewrite the equation, which we reproduce for convenience:

$$y_i = B_1 \frac{x_{1i} z_{2i}}{z_{1i}} + B_2 x_{2i} z_{2i} + B_3 z_{1i} z_{2i} + e_i z_{2i} \qquad (4')$$

Rewriting this equation reveals two important points. First, the first derivative with respect to the independent variable (x_2) is $B_2 z_2$. Once again, some of the assumptions of OLS are violated. For instance, like Study 2 this equation includes several product terms, but the equation contains none of the associated lower-order terms. Second, like Study 2, the error term is now multiplied by z_2. Because z_2 is a component of every independent variable, control variable, and the error term, endogeneity is introduced (Semadeni et al., 2014). Our simulations will provide insight regarding the implications of these potential problems.

Results. Figures 10 and 11 display results when both a control variable and the dependent variable are ratios. Figure 10 shows B_{Med} for x_2 fluctuates drastically as the coefficient of variation of z_1 increases. In fact, when the coefficient of variation of z_1 approaches 2.5, the B_{Med} for B_2 becomes negative. Furthermore, it does not appear as though the correlation between x_1 and x_2 influences this relationship. Figure 11 displays a more consistent pattern than does Figure 10. *PerSig* for B_2 appears to decrease as the coefficient of variation of z_1 increases. This pattern does not change as a function of the correlation between x_1 and x_2.
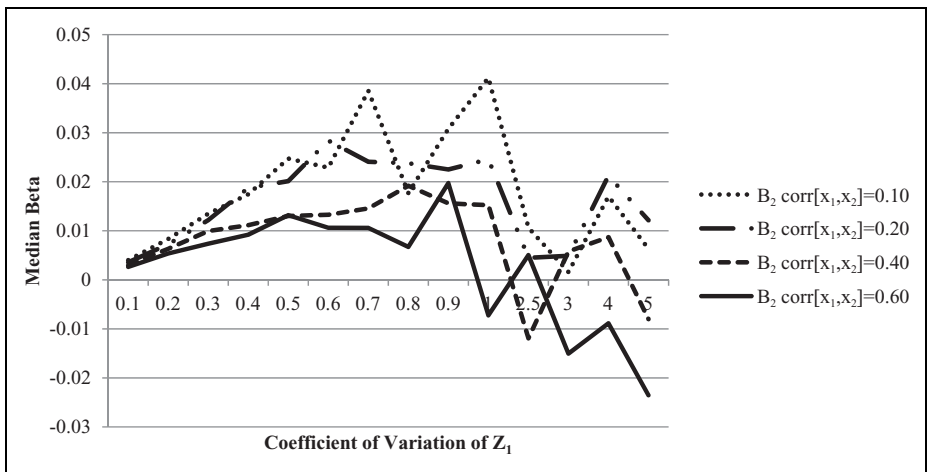
Figure 10. Median Beta When the Control and Dependent Variables Are Ratios (Study 6).
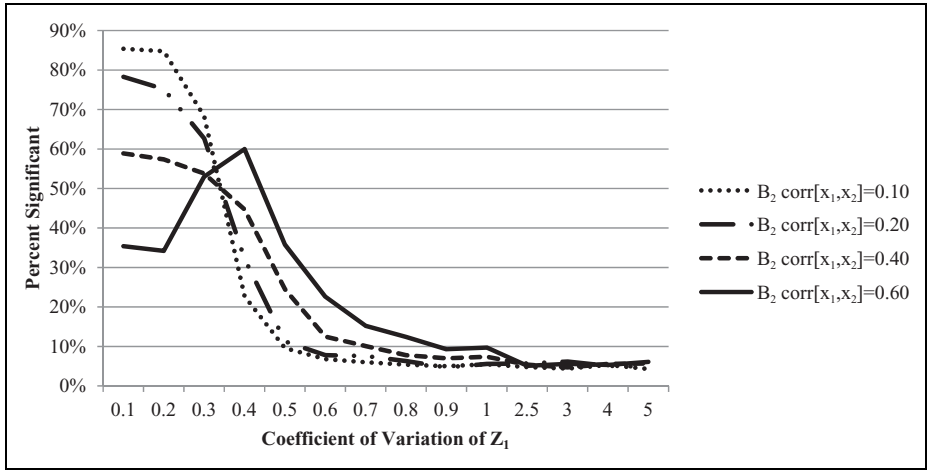

Figure 11. Percentage Significant When the Control and Dependent Variables Are Ratios (Study 6).

Discussion, Implications, and Conclusion

Discussion of the Study Results

We used simulations to illustrate how ratios in regressions introduce complexities that make it difficult to interpret the results associated with theorized relationships. In our baseline study (Study 1), the coefficient estimates and statistical power remained stable as we varied the dispersion of the scaling variable. Each of our subsequent studies, though, illustrates how ratios create inconsistencies in our results. We summarize our findings in Table 2.

Ratios as dependent variables. Study 2, in which we examined the effect of using a ratio as a dependent variable, reveals a pattern of results that differ greatly from our baseline study (Study 1). Figure 2 illustrates that the effect size of the independent variable varies wildly as the dispersion of the dependent variable's denominator increases. At the same time, the percentage of cases in which the model reports a statistically significant relationship nears zero as the dispersion increases (Figure 3). Apparently, as the dispersion of the scaling variable increases, so does the level of Type II errors.

Table 2. Summary of Simulation Results.

Study	Key Findings
Study 1: No Ratios	Effect size and percentage significant remain consistent across all CV conditions.
Study 2: Ratios as DVs	The effect size fluctuates as the CV of the denominator changes.
	Statistical significance approaches zero as the CV of the denominator increases.
Study 3: Ratios as IVs	The effect size decreases as the CV of the denominator increases.
	Statistical significance approaches zero as the CV of the denominator increases.
Study 4: Ratios as DV and IV	Effect size and statistical significance vary based on the correlation between the denominators of the ratios.
	Generally, the effect size increases as the CV of the denominator increases.
	Generally, statistical significance increase as the CV of the denominator increases.
Study 5: Ratios as Control Variables	As the CV of the denominator of the ratio control variable increases, the effect size and statistical significance of the non-ratio predictor increase.
	These effects depend on the correlation between the non-ratio predictor and the numerator of the ratio.
Study 6: Ratios as DVs and Control Variables	Effect size and statistical significance of the non-ratio variable vary dramatically as the CV of the denominator of the ratio control variable increases.
	The statistical significance of the non-ratio predictor approaches zero as the CV of the denominator of the ratio increases.

Note: These simulations alter only the CVs of the denominators of ratios. The inconsistent results reported herein become even more pronounced when also altering the CVs of the numerators of the ratios.

Stated more plainly, using ratios as dependent variables can result in an almost complete loss of statistical power.

Ratios as independent variables. While the results in Study 2 reveal unstable parameter estimates for the independent variable when the dependent variable is a ratio, the results of Study 3 show a clearer pattern when the independent variable is a ratio. Specifically, Figure 4 shows a clear decrease in the median beta estimate for the independent variable as the dispersion of the denominator increases. Similar to Study 2, Figure 5 shows the percentage of statistically significant relationships nears zero as the dispersion in the denominator increases. Study 3 demonstrates that when using ratios as independent variables, it becomes increasingly unlikely to find statistically significant effects as the dispersion of the denominator increases. Stated differently, in such conditions the likelihood of making Type II errors increases dramatically.

Ratios as both dependent and independent variables. Studies 2 and 3 examine the use of ratios as either a dependent or independent variable only, but Study 4 examines ratios as both dependent and independent variables. In this context, we find that the model results in increasing beta coefficients and statistical significance as the dispersion of the denominator increases (even when using different denominators to scale the independent and dependent variables). As shown in Figure 7, the statistical significance is overestimated. As the coefficient of variation increases, models are more likely to report relationships between two variables even when such a relationship does not exist (i.e., Type I error).

Ratios as control variables. In Study 5 we shift our focus to examine the impact of using a ratio on the results of another non-ratio predictor, which often occurs when a ratio is used as a control variable. Figures 8 and 9 suggest that as the dispersion of the denominator of a ratio control variable increases, the parameter estimate and statistical significance of the non-ratio predictor increases. Moreover, we find that these inconsistencies become more dramatic as the correlation between the hypothesized variable and the numerator of the control variable increase. Stated differently, the use

of ratios in a multiple regression model can increase the potential for Type I errors in the non-ratio predictors, and this increase is further amplified as the correlation between the control and the numerator of the independent variable increases.

Ratios as both control and dependent variables. Our final study examined the use of ratios as both control and dependent variables. Figures 10 and 11 suggest that, once again, the dispersion of the denominator influences the coefficient estimates and statistical significance of the non-ratio independent variable. In fact, the results of this study resemble those of Study 2, which examined only dependent variables as ratios. In both cases the coefficient estimates are very erratic, and statistical power approaches zero.

Rationale for the Findings in Our Studies

The findings in our studies are not anomalies. Indeed, research on empirical estimation can provide insight to the rationale and intuition for why results change when using ratios in the model (e.g., Cohen et al., 2003; Edwards et al., 2009; Kennedy, 2008). As we highlighted in the discussions of the studies, the use of ratios may violate the assumptions of OLS regression. Put differently, the mathematical derivations of the marginal effect of x_1 on y are insufficient to predict parameter estimates when the assumptions of OLS regression are violated.

Although problems created by ratio variables have been studied for over a century, we offer a number of insights regarding implications to regression in contexts of interest to strategic management scholars. For example, Pearson (1896) provided a formula to estimate the correlations between two variables (e.g., y, x) that are scaled by other variables (e.g., z_1 and z_2) to create ratios.

$$r_{\left(\frac{y}{z_1}, \frac{x}{z_2}\right)} = \frac{r_{xy} V_y V_x - r_{yz_2} V_y V_{z_2} - r_{xz_1} V_x V_{z_1} + r_{z_1 z_2} V_{z_1} V_{z_2}}{\sqrt{V_y^2 + V_{z_1}^2 - 2r_{yz_1} V_y V_{z_1}} \sqrt{V_x^2 + V_{z_2}^2 - 2r_{xz_2} V_x V_{z_2}}} \tag{5}$$

Our adaptation of this formula, shown in Equation 5, illustrates that the correlation between two variables changes as one (or both) of those variables is scaled. Stated differently, Pearson (1896) showed that two scaled variables can become more or less correlated as the coefficient of variation of the denominator(s) changes. However, we note that the predicted correlations from this formula are not *exactly* what we find in our simulations. We have two explanations.

First, our simulations involve OLS regressions and not simple correlations. As we illustrated in each study, rewriting the regression equations unmasks complex product terms that are impossible to interpret given the way the models are specified. Indeed, the problems of interpreting product terms that carry interactions without the necessary main effects are well-known (Edwards et al., 2009; Kennedy, 2008). In addition, rewriting equations illustrates the endogeneity that occurs when the dependent variable is a ratio. We show how the denominator becomes a component of every independent variable as well as the error term.

Second, Pearson's formula originally included higher ordered terms, but he dropped them because he assumed the coefficients of variation (e.g., v_y, v_x, etc.) of the variables would remain small (Dunlap et al., 1997; Pearson, 1896). Our simulated variables mimic popular data often published in strategy research, however, and include coefficients of variation exceeding one.

Implications for Research in Strategic Management

Our simulations offer a number of implications for strategy researchers, but we would also point out that our work has implications for scholars in other fields and areas of management—particularly those that use ratios consisting of variables with high coefficient of variation values. Our results

indicate that the dispersion of the denominators can change results, even when the underlying relationships among the unscaled variables remained constant. We also underscore that the effect of ratios extends beyond problems interpreting the effects of the ratio variable itself. Specifically, our results show that using a ratio as a control variable can cause the effect of a hypothesized variable on a dependent variable to change, even when the correlation between the dependent and independent variable of interest remains constant.

It is also important to highlight that our results apply to a wide array of ratios that might be used in strategy research. Given the prominent role of firm size in strategy research, the examples throughout our simulations described firm size as the scaling variable. However, our results apply to all ratios—even those that do not use firm size as the denominator (e.g., proportion of outside directors, price-to-earnings). This point is important to stress, as our review of the literature in Table 1 documents that dispersion varies dramatically across different variables, samples, and transformations.

Although we focus on simulations and empirical results, our studies have important implications for theoretical development. Our results show that relationships involving ratios can change based solely on changes in the coefficients of variation of the variables used in the ratios. This may have a significant impact on replicability and the accumulation of a systematic body knowledge, which is an issue of increasing importance to strategy researchers (Bettis, 2012; Hubbard, Vetter, & Little, 1998). The use of tools to gauge generalizability, such as meta-analyses, for example, would reveal misleading population effects and variability in the effect sizes that would be impossible to explain substantively (Dalton, Daily, Certo, & Roengpitya, 2003; Dalton, Daily, Ellstrand, & Johnson, 1998; King, Dalton, Daily, & Covin, 2004). Nevertheless, a review of these same problems in other disciplines shows that researchers are generally reluctant to believe and/or accept that ratios may impede advances in theoretical understanding. In the field of biology, for instance, Atchley and Anderson (1978, p. 71) state that "any criticism of the use of ratios by biologists elicits a Pavlovian response in some circles similar to that resulting from criticisms of motherhood, America, and apple pie." Indeed, we acknowledge the difficulty of changing a paradigm regarding how some of the most fundamental constructs in a discipline are operationalized and modeled. Fortunately, there are a number of straightforward and easy to implement solutions to address this issue.

Recommendations for Strategy Scholars

Use unscaled variables/control for scale. Perhaps the most important recommendation involves focusing on the numerators of the ratios that are used so often in strategy research. We believe strategy research would benefit greatly from focusing on variables such as net income, R&D expenditures, and other strategic variables, while using firm size or other factors that reflect a firm's scale as separate control variables. Given the inconsistencies that result from analyses involving ratios, a renewed focus on these numerators would allow for more robust findings and successful replications. Scholars would also get a clearer understanding of the effect of the scaling variables themselves. Although raw or unscaled variables in strategy research may introduce heteroscedasticity, most statistical software packages (e.g., Stata, SAS, etc.) allow researchers to account for this property using robust standard errors.

We are also hopeful that by focusing on the numerator, scholars will think more about their theoretical constructs of interest. Variables such as ROA and R&D intensity have become central in strategy. At the same time, though, customers purchase goods and services with currency, not units of ROA. Furthermore, companies pay the salaries of employees with currency and not in percentages of sales. Thinking in these terms may help strategy scholars strengthen the connection between the theoretical construct of interest and the measure used to operationalize it.

Interactions. We agree with the contention that that interactions can replace ratios (Kronmal, 1993). The use of a scaling variable as a denominator in a ratio represents a somewhat complicated construct. For instance, a researcher might hypothesize that R&D intensity (R&D spending divided by assets) is positively associated with firm performance. This operationalization, however, actually indicates that the relationship between investment in R&D and firm performance depends on firm size. If this truly represents the hypothesis of interest, the appropriate test involves an explicit interaction between R&D and firm size (e.g., R&D spending * assets). This would allow researchers to understand the main effects of R&D and sales on net income, and the interaction would help researchers understand how the relationship between R&D and net income changes as firm size changes.

Using ratios. Despite the problems with ratios, we fully understand that some researchers will insist on using them in analyses. Scholars in other fields recognize the distinction between controlling vs. scaling for a variable denoting size (e.g., Smith, 2005). Our simulations show that using unscaled variables is the most effective technique to control for a denominator of interest. When using a ratio as an independent variable, however, it is imperative to include the components of the ratio (i.e., lower order terms) as covariates for proper interpretation (Kennedy, 2008; Kronmal, 1993). If scholars still insist on using ratios, we suggest that, at the very least, they report results of models with unscaled data for purposes of comparison and full disclosure.

It is perhaps most important to report unscaled variables when the researcher has hypothesized an effect of a non-ratio predictor in a model that also includes a control variable operationalized as a ratio. As Study 5 illustrates, the effect of the hypothesized predictor and chances of confirming the hypothesis are inflated as the coefficient of variation increases. In fact, confirming a hypothesis about a non-ratio predictor in a model that also includes a substantive predictor may be virtually guaranteed once the coefficient of variation of the scaling variable reaches a certain level (approximately 0.5 in our simulation).

We should also note that some scholars may suggest that the construct relating the numerator to the denominator is of interest. Research highlights, though, that such ratios only properly control for the denominator when the two variables are strictly proportional (Barnes, 1987; Lev & Sunder, 1979; Sollberger & Ehlert, 2016). Stated differently, strict proportionality exists when the relationship between the numerator and denominator is a straight line that passes through the origin (Sollberger & Ehlert, 2016). When there is error in this relationship or this relationship does not pass through the origin, this assumption is necessarily violated. We cannot think of any commonly used ratios in strategy research that fulfill these criteria.

Wiseman (2009) provided another alternative for scholars considering the use of ratios in their statistical models. In particular, he highlighted how scholars can perform a two-stage process that essentially controls for the effects of a scaling variable without using ratios. The first stage of this procedure is a regression where x_1 is the dependent variable and z_1 is the independent variable. Scholars can then store the residuals from the first stage regression, which represent the variance in x_1 not attributable to z_1. The second stage of the regression includes these residuals as a substitute independent variable (i.e., instead of x_1 as the independent variable). The intent of such an approach is to capture variance in the independent variable unrelated to the scaling variable, which provides the benefits of creating a ratio without the detriments (Wiseman, 2009). In supplementary analyses we found that this technique is inferior to simply using unscaled variables, both in bias and efficiency.

Replications. In light of our findings, our final suggestion is that the field of strategy needs a series of replications examining important relationships that involve constructs typically expressed as ratios. As we described above, parameter estimates as well as statistical power are a function of the coefficient of variation of the components of the ratios in a regression model, and thus, it is

unclear that results from strategy research involving ratios can be interpreted unequivocally. Replications of empirical models that feature ratios would be valuable to assess the validity of our accumulated knowledge. A good example of this is a study by Bromiley, Rau, and Zhang (2017), which contrasted the effects of R&D expenditures (i.e., unscaled independent variable) and R&D intensity (i.e., ratio independent variable) on firm risk. Perhaps unsurprisingly, they found very different results depending on whether or not R&D was operationalized as a ratio. Our simulations suggest this phenomenon perhaps generalizes to many more constructs beyond R&D. Although theoretical novelty and interestingness are criteria that are often considered to gauge the contribution of a management article, progress in the field of strategy requires replications. It is our hope that editors of the journals that publish articles on strategic management will appreciate the importance of this type of contribution.

Conclusion

Although this article may be quite distressing, there are several reasons why we are optimistic that scholars in management and strategy can and will adapt. First, the alternatives to using ratios we outlined above are remarkably easy to implement. No new training is required. Second, many of our simulations suggest that using unscaled data results in greater statistical power. In many cases, scholars are more likely to find significant effects if they exist when using unscaled data. Third, alternatives to ratios can provide new substantive insights that are crystal clear and practically important. As an example, the implications of results indicating how the level of R&D expenditures influences net income for different sized firms (explicit consideration of interactions between the substantive variable of interest and the scaling variable) are much clearer than results indicating there is an effect of R&D/sales on net income/assets for firms with average levels of assets and R&D expenditures. Finally, using unscaled data should result in findings that are more consistent across studies. This consistency, in turn, will promote the type of knowledge accumulation that will allow the field to mature and increase in practical relevance. In the end, the issues we identified in this article should be viewed not as a dark cloud hanging over the discipline of strategic management, but a bright ray of hope. Our wish is that readers will come to share this sanguine view of our findings and the opportunities that follow.

Notes

1. Our review also revealed that a majority of these ratios included measures of firm size (e.g., assets, revenues, employees, etc.) as the denominator.
2. In supplementary analyses, we also generated y as a function of the independent covariates. The results are substantively similar.
3. In a supplementary approach, we altered the standard deviation of z1 instead of the mean and the standard deviations of all covariates in proportion to ensure consistent variance explained. The results were substantively similar to those reported here.
4. It is important to note that different estimators used to address the nature of a study's data (e.g., fixed effects, random effects, two-stage models) do not yield remarkably different results than those we report in this study. In supplementary simulations, we generated nested observations to mimic the properties of panel data

(e.g., firms nested within years). We used this to examine the effects of ratios in fixed or random effects models. We also simulated noncontinuous dependent variables (such as a count) to verify ratios in Poisson, negative binomial, and even probit/logit models produce similar problems.

5. In supplementary analyses, we did not include assets in the model. Including or excluding the scaling variable as a control does not substantively change the results. We retain it as a control for our models since it is correlated with both R&D and net income in our simulations, so excluding slightly inflates the betas and percentage significant owing to omitted variable bias. As a further check to ensure our results are not an artifact of the correlation between the independent and scaling variables, in supplementary analyses we observed similar patterns when the correlation between R&D and assets was .7.

6. In the review process, we included figures to demonstrate the stability of these parameter estimates in the baseline study. Although we omitted these figures from the final article for space reasons, they are available from the authors on request.

References

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage.

Atchley, W. R., & Anderson, D. (1978). Ratios and the statistical analysis of biological data. *Systematic Zoology*, *27*(1), 71-78.

Barnes, P. (1987). The analysis and use of financial ratios: a review article. *Journal of Business Finance & Accounting*, *14*(4), 449-461.

Bertrand, O., & Capron, L. (2015). Productivity enhancement at home via cross-border acquisitions: The roles of learning and contemporaneous domestic investments. *Strategic Management Journal*, *36*(5), 640-658.

Betschinger, M.-A. (2015). Do banks matter for the risk of a firm's investment portfolio? Evidence from foreign direct investment programs. *Strategic Management Journal*, *36*(8), 1264-1276.

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

Bosco, F. A., Aguinis, H., Singh, K., Field, J. G., & Pierce, C. A. (2015). Correlational effect size benchmarks. *Journal of Applied Psychology*, *100*(2), 431.

Brambor, T., Clark, W. R., & Golder, M. (2006). Understanding interaction models: Improving empirical analyses. *Political Analysis*, *14*, 63-82.

Bromiley, P., Rau, D., & Zhang, Y. (2017). Is R&D risky? *Strategic Management Journal*, *38*(4), 876-891.

Brown, S., Lo, K., & Lys, T. (1999). Use of R2 in accounting research: measuring changes in value relevance over the last four decades. *Journal of Accounting and Economics*, *28*(2), 83-115.

Campbell, J. T., Campbell, T. C., Sirmon, D. G., Bierman, L., & Tuggle, C. S. (2012). Shareholder influence over director nomination via proxy access: Implications for agency conflict and stakeholder value. *Strategic Management Journal*, *33*(12), 1431-1451.

Carow, K., Heron, R., & Saxton, T. (2004). Do early birds get the returns? An empirical investigation of early-mover advantages in acquisitions. *Strategic Management Journal*, *25*(6), 563-585.

Certo, S. T., Busenbark, J. R., Woo, H.-S., & Semadeni, M. (2016). Sample selection bias and Heckman models in strategic management research. *Strategic Management Journal*, *37*(1), 2639-2657.

Chadwick, C., Super, J. F., & Kwon, K. (2015). Resource orchestration in practice: CEO emphasis on SHRM, commitment-based HR systems, and firm performance. *Strategic Management Journal*, *36*(3), 360-376.

Chang, S.-J., & Shim, J. (2015). When does transitioning from family to professional management improve firm performance? *Strategic Management Journal*, *36*(9), 1297-1316.

Chen, C.-M., Delmas, M. A., & Lieberman, M. B. (2015). Production frontier methodologies and efficiency as a performance measure in strategic management research. *Strategic Management Journal*, *36*(1), 19-36.

Chen, G. (2015). Initial compensation of new CEOs hired in turnaround situations. *Strategic Management Journal*, *36*(12), 1895-1917.

Chizema, A., Liu, X., Lu, J., & Gao, L. (2015). Politically connected boards and top executive pay in Chinese listed firms. *Strategic Management Journal*, *36*(6), 890-906.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155-159.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum.

Curran-Everett, D. (2013). Explorations in statistics: The analysis of ratios and normalized data. *Advances in Physiology Education*, *37*(3), 213-219.

Dalton, D. R., Daily, C. M., Certo, S. T., & Roengpitya, R. (2003). Meta-analyses of financial performance and equity: Fusion or confusion? *Academy of Management Journal*, *46*(1), 13-26.

Dalton, D. R., Daily, C. M., Ellstrand, A. E., & Johnson, J. L. (1998). Meta-analytic reviews of board composition, leadership structure, and financial performance. *Strategic Management Journal*, *19*(3), 269-290.

Dezső, C. L., Ross, D. G., & Uribe, J. (2016). Is there an implicit quota on women in top management? A large-sample statistical analysis. *Strategic Management Journal*, *37*(1), 98-115.

Diestre, L., Rajagopalan, N., & Dutta, S. (2015). Constraints in acquiring and utilizing directors' experience: An empirical study of new-market entry in the pharmaceutical industry. *Strategic Management Journal*, *36*(3), 339-359.

Dunlap, W. P., Dietz, J., & Cortina, J. M. (1997). The spurious correlation of ratios that have common variables: A Monte Carlo examination of Pearson's formula. *Journal of General Psychology*, *124*(2), 182-193.

Edwards, J. R., Lance, C., & Vandenberg, R. (2009). *Statistical and methodological myths and urban legends: Doctrine, verity and fable in the organizational and social sciences*. New York, NY: Routledge.

Feldman, E. R. (2016). Managerial compensation and corporate spinoffs. *Strategic Management Journal*, *37*(10), 2011-2030.

Flammer, C. (2015). Does product market competition foster corporate social responsibility? Evidence from trade liberalization. *Strategic Management Journal*, *36*(10), 1469-1485.

Gupta, A., Briscoe, F., & Hambrick, D. C. (2017). Red, blue, and purple firms: Organizational political ideology and corporate social responsibility. *Strategic Management Journal*, *38*, 1018-1040. doi:10.1002/smj.2550

Harrison, D. A., & Klein, K. J. (2007). What's the difference? Diversity constructs as separation, variety, or disparity in organizations. *Academy of Management Review*, *32*(4), 1199-1228.

Hill, A. D., Upadhyay, A. D., & Beekun, R. I. (2015). Do female and ethnically diverse executives endure inequity in the CEO position or do they benefit from their minority status? An empirical examination. *Strategic Management Journal*, *36*(8), 1115-1134.

Hubbard, R., Vetter, D. E., & Little, E. L. (1998). Replication in strategic management: Scientific testing for validity, generalizability, and usefulness. *Strategic Management Journal*, *19*(3), 243-254.

Humphery-Jenner, M. (2014). Takeover defenses, innovation, and value creation: Evidence from acquisition decisions. *Strategic Management Journal*, *35*(5), 668-690.

Jasienski, M., & Bazzaz, F. A. (1999). The fallacy of ratios and the testability of models in biology. *Oikos*, *84*(2), 321-326.

Josefy, M., Kuban, S., Ireland, R. D., & Hitt, M. A. (2015). All things great and small: Organizational size, boundaries of the firm, and a changing environment. *Academy of Management Annals*, *9*(1), 715-802.

Judge, W. Q., Witt, M. A., Zattoni, A., Talaulicar, T., Chen, J. J., Lewellyn, K., . . . Van Ees, H. (2015). Corporate governance and IPO underpricing in a cross-national sample: A multilevel knowledge-based view. *Strategic Management Journal*, *36*(8), 1174-1185.

Kang, J. (2016). Labor market evaluation versus legacy conservation: What factors determine retiring CEOs' decisions about long-term investment? *Strategic Management Journal*, *37*(2), 389-405.

Kennedy, P. (2008). *A guide to econometrics* (2nd ed.). Oxford, UK: Blackwell.

Ketchen, D. J., Boyd, B. K., & Bergh, D. D. (2008). Research methodology in strategic management: Past accomplishments and future challenges. *Organizational Research Methods*, *11*(4), 643-658.

Kim, J.-H. (1999). Spurious correlation between ratios with a common divisor. *Statistics & Probability Letters*, *44*(4), 383-386.

King, D. R., Dalton, D. R., Daily, C. M., & Covin, J. G. (2004). Meta-analyses of post-acquisition performance: indications of unidentified moderators. *Strategic Management Journal*, *25*(2), 187-200.

Krause, R., Priem, R., & Love, L. (2015). Who's in charge here? Co-CEOs, power gaps, and firm performance. *Strategic Management Journal*, *36*(13), 2099-2110.

Kronmal, R. A. (1993). Spurious correlation and the fallacy of the ratio standard revisited. *Journal of the Royal Statistical Society A*, *156*(3), 379-392.

Lev, B., & Sunder, S. (1979). Methodological issues in the use of financial ratios. *Journal of Accounting and Economics*, *1*(3), 187-210.

Liermann, M., Steel, A., Rosing, M., & Guttorp, P. (2004). Random denominators and the analysis of ratio data. *Environmental and Ecological Statistics*, *11*(1), 55-71.

Lim, E. N. K. (2015). The role of reference point in CEO restricted stock and its impact on R&D intensity in high-technology firms. *Strategic Management Journal*, *36*(6), 872-889.

Makri, M., Hitt, M. A., & Lane, P. J. (2010). Complementary technologies, knowledge relatedness, and invention outcomes in high technology mergers and acquisitions. *Strategic Management Journal*, *31*(6), 602-628.

Neyman, J. (1952). *Lectures and conferences on mathematical statistics and probability* (2nd ed.). Washington, DC: U.S. Department of Agriculture.

O'Connell, V., & O'Sullivan, D. (2014). The influence of lead indicator strength on the use of nonfinancial measures in performance management: Evidence from CEO compensation schemes. *Strategic Management Journal*, *35*(6), 826-844.

Packard, G. C., & , & Boardman, T. J. (1988). The misuse of ratios, indices, and percentages in ecophysiological research. *Physiological Zoology*, *61*(1), 1-9.

Pearson, K. (1896). Mathematical contributions to the theory of evolution: On a form of spurious correlation which may arise when indices are used in the measurement of organs. *Proceedings of the Royal Society of London*, *60*(359-367), 489-498.

Semadeni, M., Withers, M. C., & Certo, S. T. (2014). The perils of endogeneity and instrumental variables in strategy research: Understanding through simulations. *Strategic Management Journal*, *35*(7), 1070-1079.

Shi, W., Connelly, B. L., & Sanders, W. G. (2016). Buying bad behavior: Tournament incentives and securities class action lawsuits. *Strategic Management Journal*, *37*(7), 1354-1378.

Shiu, Y. M., & Yang, S. L. (2017). Does engagement in corporate social responsibility provide strategic insurance-like effects? *Strategic Management Journal*, *38*, 455-470. doi:10.1002/smj.2494

Smith, R. (2005). Relative size versus controlling for size interpretation of ratios in research on sexual dimorphism in the human corpus callosum. *Current Anthropology*, *46*(2), 249-273.

Sollberger, S., & Ehlert, U. (2016). How to use and interpret hormone ratios. *Psychoneuroendocrinology*, *63*, 385-397.

Tyler, B. B., & Caner, T. (2016). New product introductions below aspirations, slack and R&D alliances: A behavioral perspective. *Strategic Management Journal*, *37*(5), 896-910.

Vomberg, A., Homburg, C., & Bornemann, T. (2015). Talented people and strong brands: The contribution of human capital and brand equity to firm value. *Strategic Management Journal*, *36*(13), 2122-2131.

Wiseman, R. M. (2009). On the use and misuse of ratios in strategic management research. In D. Ketchen & D. Bergh (Eds.), *Research methodology in strategy and management* (pp. 75-110). Bingley, UK: Emerald.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. Cambridge, MA: MIT press.

Wowak, A. J., Mannor, M. J., & Wowak, K. D. (2015). Throwing caution to the wind: The effect of CEO stock option pay on the incidence of product safety problems. *Strategic Management Journal*, *36*(7), 1082-1092.

Zhao, X., & Murrell, A. J. (2016). Revisiting the corporate social performance-financial performance link: A replication of Waddock and Graves. *Strategic Management Journal*, *37*(11), 2378-2388.

Zhu, H., & Yoshikawa, T. (2016). Contingent value of director identification: The role of government directors in monitoring and resource provision in an emerging economy. *Strategic Management Journal*, *37*(8), 1787-1807.

Author Biographies

S. Trevis Certo (PhD, Indiana University) is the Jerry B. and Mary Anne Chapman Professor of Business in the Department of Management and Entrepreneurship in the W. P. Carey School of Business at Arizona State University. His research focuses on corporate governance, top management teams, initial public offerings (IPOs), and research methodology.

John R. Busenbark (PhD, Arizona State University) is an assistant professor of management in the Terry College of Business at the University of Georgia. His research interests include corporate governance, information economics, capital markets, and research methodology.

Matias Kalm is a doctoral student in the W. P. Carey School of Business at Arizona State University. His research interests include corporate governance, top management teams, entrepreneurship, and research methodology.

Jeffery A. LePine (PhD, Michigan State University) is a professor and PetSmart Chair in Leadership with W. P. Carey School of Business. His research interests center on three broad areas of organizational behavior: (a) the functioning and effectiveness of small groups and teams, (b) job performance behaviors involving citizenship, voice, and adaptability, and (c) the influence of engagement and stress on employee effectiveness and well-being. He is a fellow of the American Psychological Association and the Society for Industrial and Organizational Psychology.