

# Journal of Management

<http://jom.sagepub.com>

---

## Using Competition to Measure Relatedness

Lasse B. Lien and Peter G. Klein

*Journal of Management* 2009; 35; 1078 originally published online Dec 22, 2008;  
DOI: 10.1177/0149206308328505

The online version of this article can be found at:  
<http://jom.sagepub.com/cgi/content/abstract/35/4/1078>

---

Published by:



<http://www.sagepublications.com>

On behalf of:



[Southern Management Association](#)

**Additional services and information for *Journal of Management* can be found at:**

**Email Alerts:** <http://jom.sagepub.com/cgi/alerts>

**Subscriptions:** <http://jom.sagepub.com/subscriptions>

**Reprints:** <http://www.sagepub.com/journalsReprints.nav>

**Permissions:** <http://www.sagepub.com/journalsPermissions.nav>

**Citations** <http://jom.sagepub.com/cgi/content/refs/35/4/1078>

# Using Competition to Measure Relatedness<sup>†</sup>

Lasse B. Lien\*

*Department of Strategy and Management, Norwegian School of Economics and Business Administration, Breiviksveien 40, N-5045 Bergen, Norway*

Peter G. Klein

*Department of Agricultural Economics, University of Missouri  
135 Mumford Hall, Columbia, MO 65211*

*Department of Strategy and Management, Norwegian School of Economics and Business Administration, Breiviksveien 40, N-5045 Bergen, Norway*

---

*The conventional approach to measuring interindustry relatedness uses the Standard Industrial Classification (SIC) system to capture the “distance” between industries. Although relatedness measures based on SIC codes (or equivalent classifications) are readily available and easy to compute, they do not screen effectively for the conditions under which related diversification creates value. This article constructs an alternative, survivor-based measure of interindustry relatedness and compares it to similar measures based on distances between SIC codes. The authors find that survivor-based measures consistently outperform SIC-based measures in predicting firms’ decisions to enter new markets, even when herding tendencies and motives related to mutual forbearance are taken into account.*

**Keywords:** *diversification; relatedness; measurement; corporate strategy*

---

Interindustry relatedness is one of the core concepts of corporate strategy. Relatedness is thought to affect performance (Rumelt, 1974, 1982), the direction of diversification (Chatterjee & Wernerfelt, 1991; Montgomery & Hariharan, 1991), entry mode (Chatterjee,

---

<sup>†</sup>We thank Jay Barney, Erik Døving, Nicolai Foss, Olav Kvitastein, Keld Laursen, Rich Makadok, Christine Meyer, Associate Editor Manuel Becerra, three anonymous referees, and participants at the 2005 Academy of Management Annual Meeting and the 2006 Copenhagen Conference on Strategic Management for helpful comments and suggestions. The usual caveat applies.

\*Corresponding author. Tel.: +47 55 959 726; fax: +47 55 959 780.

E-mail address: [lasse.lien@nhh.no](mailto:lasse.lien@nhh.no)

Journal of Management, Vol. 35 No. 4, August 2009 1078-1107

DOI: 10.1177/0149206308328505

© 2009 Southern Management Association. All rights reserved.

1990; Yip, 1982), organizational arrangements (Hill & Hoskisson, 1987; Markides & Williamson, 1996), financing (Kochar & Hitt, 1998; Myers & Majluf, 1984), and more. Overall, however, the empirical evidence on the effects of relatedness is mixed (see the descriptions in Hoskisson & Hitt, 1990; Lemelin, 1982; Markides & Williamson, 1994, 1996; Palich, Cardinal, & Miller, 2000; Reed & Luffman, 1986; Robins & Wiersema, 1995, 2003). The literature's only consistent finding is that diversification patterns are not random and that diversification is most likely to occur along a related path (Chatterjee & Wernerfelt, 1991; Montgomery & Hariharan, 1991; Silverman 1999). The fundamental strategic question—whether relatedness improves performance—remains unsettled.

One challenge in studying relatedness is that the decision to diversify is endogenous; the same factors that cause firms to diversify may also drive other behaviors and performance, making it difficult to disentangle cause and effect. Many recent articles in the diversification discount literature attempt to grapple with this kind of endogeneity by using self-selection and instrumental variables models and by examining the characteristics of diversified firms before they became diversified (Campa & Kedia, 2002; Chevalier, 2000; Graham, Lemmon, & Wolf, 2002; Miller, 2004, 2006; Villalonga, 2004). However, research on relatedness faces an additional difficulty: Relatedness itself is hard to conceptualize, parameterize, and measure consistently across industries. A variety of continuous and categorical measures, typically based on distances between Standard Industrial Classification (SIC) codes, have been used, but not consistently, and the measurement problem may explain the inconsistent findings on the performance effects of relatedness (Caves, Porter, & Spence, 1980; Chatterjee & Blocher, 1992; Fan & Lang, 2000; Farjoun, 1994; Hall & St. John, 1994; Hoskisson, Hitt, Johnson, & Moesel, 1993; Jacquemin & Berry, 1979; Markides & Williamson, 1994, 1996; Montgomery & Hariharan, 1991; Robins & Wiersma, 1995, 2003).

Here we describe and validate a fundamentally different approach to measuring relatedness. Our approach lets the competitive process and the knowledge of local decision makers replace the judgment of the researcher (or the SIC system) in determining what is related to what. This survivor-based (henceforth, SB) approach assumes that industries that are frequently combined within real firms in competitive markets are more closely related than industries that are rarely combined. Specifically, we measure the relatedness between a pair of industries by comparing how often they are actually combined to what one would expect if diversification patterns were random. Industries are related when this difference is large and positive, and they are unrelated if it is negative. This concept was originally suggested by Teece, Rumelt, Dosi, and Winter (1994), who used it to illustrate that coherence (nonrandomness) is a salient attribute of the diversification patterns of U.S. firms. To our knowledge, this procedure has not yet been evaluated as a method of measuring interindustry relatedness.

We argue that the SB approach is more satisfactory than the conventional approaches based on SIC distances, both theoretically and empirically. As we explain below, theory suggests that related diversification creates efficiencies only if specific theoretical conditions are met, and these conditions are difficult to measure directly. The conventional approaches, and some of the newer ones, are largely insensitive to these conditions. The SB approach, based on revealed preferences, takes them into account. To evaluate the effectiveness of the SB approach, we show, empirically, that it outperforms the conventional approaches in predicting entry decisions by diversified firms. As noted above, the direction of diversification is the one dependent variable

where relatedness has proved to yield consistent empirical support. We show that that the SB approach captures this relationship well. Of course, entry decisions reflect expected, not realized, efficiencies, so we look next at the postentry performance of firms that enter related industries. We show that the SB approach predicts the probability that a diversified firm is able to survive in a newly entered industry better than the conventional approaches do.

Finally, we check to see if SB measures of relatedness are contaminated by motives for diversification unrelated to efficiency, such as herd behavior and mutual forbearance. If entry reflects herding, as in the “information cascade” models of Banerjee (1992) and Bikhchandani, Hirshleifer, and Welch (1992), rather than independent assessments of the characteristics of particular industry combinations, then observed combinations tell us little about efficiency. However, the SB approach is a good predictor of postentry survival, which is inconsistent with a herding-based explanation for diversification. The mutual forbearance hypothesis suggests that contact between firms across markets induces implicit or explicit agreements to refrain from aggressive competition (Edwards, 1955; Gimeno, 1999; Greve & Baum, 2001). SB measures of relatedness are particularly prone to incorporate such behaviors because they are built by counting the frequencies of multimarket contact across industries. We show, however, that the SB measure is even more effective at predicting entry into fragmented industries, in which forbearance motives are unlikely to play a role, than into concentrated markets. In short, the SB approach appears to capture, effectively, the kinds of relatedness that matter for efficiency.

The discussion is organized as follows: We begin by reviewing the resource-based approach to relatedness, explaining the conditions under which diversification should improve efficiency and summarizing existing measures of relatedness. Next, we describe the SB approach and explain how it can alleviate the shortcomings of existing measures. We turn next to the empirical work, comparing the quality of SB and SIC-based measures in predicting market-entry decisions by diversifying firms, along with the degree to which entry decisions are reversed. Finally, we show that our analysis is robust to possible contamination from herding or mutual forbearance.

## **Diversification, Relatedness, and Efficiency**

Diversification decisions figure prominently in both the strategic positioning and resource-based approaches to competitive and corporate strategy. Both approaches emphasize the potential efficiency gains from exploiting scope economies, although Porter (1980) pays particular attention to the stand-alone attractiveness of the target industry and the costs of entry. Transaction cost economics, following Williamson (1975), has emphasized the potential for diversified acquisitions to generate efficiencies in the internal capital market. Alchian (1969) and Williamson (1975), and more recently Gertner, Scharfstein, and Stein (1994) and Stein (1997), argue that internal capital markets have advantages where access to external funds is limited. The central office of the diversified firm can use informational advantages, residual control rights, and its ability to intervene selectively in divisional affairs to allocate resources within the firm better than the external capital markets would do if the divisions were stand-alone firms. However, other writers argue that cross-subsidization is harmful, leading to rent seeking by divisional managers (Scharfstein & Stein, 2000), bargaining problems within the

firm (Rajan, Servaes, & Zingales, 2000), or bureaucratic rigidity (Shin & Stulz, 1998). More generally, managers may diversify, especially by acquisition, primarily to increase their compensation, job security, or span of control (Amihud & Lev, 1981).

The SB approach focuses on which combinations of industries are systematically preferred by managers, although it is agnostic about the precise mechanisms by which such combinations are chosen. However, as we indicate here, the pattern of combinations we observe cannot be explained by herding or mutual forbearance, suggesting that efficiency, not agency or cognitive bias, drives these patterns. The resource-based view suggests two possible sources of efficiency, substitutability among resources and complementarities across resources, all in the presence of positive transaction costs. Advantages of the internal capital market could also play a role, although this approach has generally been applied to unrelated, rather than related, diversification.

Efficiency explanations for related diversification typically center on economies of scope. Scope economies justify related diversification when three conditions are met. First, resources must be functional substitutes across industries, meaning that resources used in one industry can be used in another. Second, these resources must be at least partially indivisible. Consider a situation in which a resource in industry A is a perfect substitute for a resource in industry B. If resources are perfectly divisible, then a firm realizes no benefit from diversifying from industry A into industry B. If resources are not perfectly divisible,<sup>1</sup> single-business firms and unrelated diversifiers are left with costly excess capacity. This insight goes back to Penrose (1959), who noted that in the course of their normal operations firms continuously generate new resources and excess capacity in existing resources. If these resources are not fully exploited in the firm's existing businesses, they are potentially deployable at low marginal cost in a new business. Accordingly, relatedness is valuable only when some indivisibilities exist.

Third, there must be transaction costs in the market for excess capacity. As Teece (1980, 1982) points out, although indivisibilities imply joint production, they do not explain why joint production must be organized within a single firm. If the excess capacity created by indivisibilities can be traded in well-functioning markets, then single-business firms and unrelated diversifiers can simply sell or rent out their excess capacity or buy exactly the capacity they need from others. In other words, absent transactional difficulties, two separate firms could contract to share the inputs, facilities, or whatever accounts for the relevant scope economies. If they do not, it must be because the costs of writing or enforcing such a contract are greater than the benefits from joint production. Whether the firms will integrate thus depends on the comparative costs and benefits of contracting, not on the underlying production technology. Indeed, if contracting costs are low, the related diversifier may actually compete at a disadvantage relative to the single-business firm, because the diversified firm faces the additional bureaucratic costs of low-powered incentives, increased complexity, and so on (Williamson, 1985).

More recent attention has focused not on resource substitutability but on resource complementarity (Christensen & Foss, 1997; Foss & Christensen, 2001; Larsson & Finkelstein, 1999; Teece et al., 1994). Complementarities exist when investment in one industry increases the value of resources used in another industry or when decisions about resource use in one industry affect similar decisions in another. These positive spillovers create a quantitative

and qualitative coordination problem that may be best managed within a diversified firm (Milgrom & Roberts, 1992; Richardson, 1972).<sup>2</sup> To explain why this coordination problem cannot be solved in the market (i.e., between single-business firms or unrelated diversifiers), we must again appeal to some form of contracting costs. Hence, transaction costs are also relevant to situations involving complements.<sup>3</sup>

In sum, similarity among a diversified firm's industries should not be expected to create efficiency advantages unless accompanied by indivisibilities or positive spillovers and, in either case, positive transaction costs. Without the prospect of efficiency advantages, relatedness is not likely to matter (or must matter for other reasons than we usually assume). Unfortunately, as discussed below, developing measures of relatedness that screen for these conditions in a convincing manner is extremely difficult.

Of course, diversification may also reflect agency problems or attempts to exploit market power, rather than efficiency. We deal with these first by controlling for firm- and industry-specific factors associated with agency and market power when validating the SB measure and second with our analyses of herding and mutual forbearance. Although these do not exhaust the possible set of nonefficiency explanations for related diversification, our analysis suggests that SB measures do capture key efficiency aspects of relatedness.

## Existing Measures of Relatedness

Measures of relatedness within strategic management research can be divided into three major categories: categorical measures, continuous SIC-based measures, and a few recent developments. As argued below, none of these effectively captures the conditions described in the previous section, the conditions under which relatedness generates value through efficiency gains.

### *Categorical Measures*

The categorical approach, associated most closely with Rumelt (1974), is perhaps the best-known attempt to capture relatedness. Based on three ratios, Rumelt classified diversification strategies into four broad categories (nine if subcategories are included): single-business firms, dominant-business firms, related firms, and unrelated firms. The ratios used for classification are as follows: The *specialization ratio* is the proportion of a firm's revenue attributable to its largest single business. The *related ratio* is the proportion of a firm's revenue attributable to its largest group of related businesses. The *vertical ratio* is the proportion of a firm's revenue arising from all byproducts, intermediate products, and end products of a vertically integrated sequence of processing activities. The classification of businesses into related and unrelated—needed to compute the related ratio—is done subjectively, using similarities in inputs, production technology, distribution channels, and customers. As others have pointed out, there are potential problems with this procedure (Markides & Williamson, 1996; Robins & Wiersema, 1995). Besides the reliability of the subjective element in the classifications, this procedure measures relatedness on a nominal level, only allowing comparisons within group averages.

Moreover, this procedure mainly captures the degree to which resources are potential substitutes across industry boundaries, which is only the case when there are large similarities in inputs, production technology, distribution channels, and customers. These measures do not capture indivisibilities, which determine whether excess capacity is likely to develop, nor do they incorporate any notion of transaction costs impeding coordination between firms. Because both indivisibilities and transaction are necessary for economies of scope to benefit related diversifiers, these measures are prone to exaggerate relatedness in some instances (i.e., where resources are close substitutes, but these additional conditions are not met). The implicit focus on similarities and economies of scope also raises the concern that such a procedure may not capture complementarity well, which implies that it will underestimate relatedness in other instances (Foss & Christensen, 2001).

### *Continuous SIC-Based Measures*

The use of continuous SIC-based measures is by far the most popular approach in the strategy literature (Robins & Wiersema, 2003). These measures include the entropy index (Jacquemin & Berry, 1979) and the concentric index (Caves et al., 1980). Compared to Rumelt's categorical measures, these measures are less dependent on subjective assessments about the degree to which particular industries are related. The assignment of SIC codes to industries by the Census Bureau does of course involve some subjectivity, but the subjective component is consistent across studies that use SIC codes to compute relatedness. SIC-based measures also allow relatedness to be measured in intervals. The 2-, 3-, and 4-digit levels in the SIC system are treated as points on an underlying scale of relatedness, and arithmetic values are assigned to the distances.

However, the use of distances in the SIC system also introduces problems. Hall and St. John (1994), for example, found that categorical and continuous measures capture different aspects of relatedness and even question whether they capture the same underlying construct. Continuous measures are built on an assumption that industries are homogeneous within category levels, which is problematic if the breadth of the industry classifications vary, which most observers agree that they do (Robins & Wiersema, 1995; Rumelt, 1982). It also assumes that industries equally distant within the SIC hierarchy are equally dissimilar, which is problematic. Most important for present purposes, however, SIC-based measurements are no better than categorical measures in capturing indivisibilities and transaction costs.<sup>4</sup> In other words, even if distance among SIC codes is a good proxy for resource substitutability, these measures will tend to exaggerate relatedness. Foss and Christensen (2001) also point out that SIC-based procedures have an implicit bias toward economies of scope and thus are unlikely to capture complementarity, suggesting that this type of relatedness is prone to be underestimated.

### *Recent Developments*

One way to measure relatedness more consistently with the theoretical literature is to focus on resources that are particularly likely to generate excess capacity and positive spillover and

be subject to positive transaction costs. If such types of resources can be identified, similarities among such resources should be more likely to enhance efficiency. One particularly promising approach is the use of data from patent filings to mark technology flows between industries, which in turn indicates the extent to which technological resources in one industry are valuable in another (Breschi, Lissoni, & Malerba, 2003; Engelsman & van Raan, 1992; Jaffe, 1986; Kim & Kogut, 1996; Laursen & Melicani, 2000; Robins & Wiersema, 1995; Silverman, 1999; Piscitello, 2000).<sup>5</sup> The results of these studies are generally in line with theoretical predictions from the relatedness literature, more so than studies using categorical and SIC-based measures, perhaps because relatedness is measured more appropriately. Indeed, technological resources are likely to be imperfectly divisible; some technological resources, like patents, can be described as quasi-public goods, meaning that their use in one business does not preclude their use in another (if the technology or knowledge can be replicated at zero or low marginal cost). Such technology flows may also indicate the extent to which there are dynamic complementarities across industries. Finally, there are reasons to believe that trading in technological resources is often subject to high transaction costs (Teece, 1986). Tacit knowledge may make it difficult to transfer technological resources by contract. Alternatively, technological resources may be easy to transfer, but the value they represent may be difficult to appropriate once the knowledge has been revealed to a potential buyer.

However, such measures are not without limitations. First, they can capture only relatedness associated with technological resources, and can do so in patent-intensive industries. This indicates some severe restrictions on where these measures can be applied and the forms of relatedness they can capture. Moreover, these measures are probably noisy even under favorable circumstances. Not all technological resources are quasi-public goods; some, for example, may be linked to knowledge that cannot be easily replicated. This can be the case if technology transfer requires knowledge that cannot be separated from one or a group of individuals (perhaps due to its tacit character). The capacity of these individuals may be exhausted in existing applications, leaving the condition of excess capacity unsatisfied. Also, there are several types of technological resources where market trading is indeed feasible (Levin, Klevorick, Nelson, Winter, Gilbert, & Griliches, 1987; Silverman, 1999; Teece, 1986). The prevalence of technology and patent licensing arrangements illustrates this point.

Other measures used occasionally in the relatedness literature include ones based on human resource profiles (Farjoun, 1994), input ratios (Montgomery & Hariharan, 1991), commodity flows (Fan & Lang, 2000), and more. Although these may vary in their abilities to capture to what extent resources in one industry are substitutes for or complements to resources in another, none offer a full solution for capturing the conditions of indivisibility, spillovers, and positive transaction costs.

More generally, measures based on patent concordances, human resource profiles, input ratios, and commodity flows capture a different kind of relatedness than do measures based on SIC codes. The former capture relatedness of resources, whereas SIC-based measures capture relatedness of products or markets. Although product-market relatedness is an important dimension of corporate strategy, it is not the kind of relatedness that theory tells us should lead to sustained competitive advantage. SB measures, as we argue below, subsume both resource and product-market relatedness as well as other diversification drivers (institutional effects, forbearance, and the like).

## A Survivor-Based Approach to Relatedness

The core of the survivor principle is that the competitive process screens for efficiency and does so well enough that a sample of competitive firms will be dominated by the decisions or behaviors that are efficient (at least in a comparative sense; Alchian, 1950: 211).<sup>6</sup> The argument relies on two key assumptions. One is that firms making negative profits will, unless some corrective measure is taken, lose resources and ultimately become extinct, whereas firms making positive profits will acquire resources and grow. The second assumption is that the desire for profit provides a strong incentive for the less successful firms to imitate the behavior of the more successful firms. Although few believe that the competitive process performs this screening perfectly,<sup>7</sup> research in industrial organization, organizational economics, and strategic management seems to indicate a fairly optimistic view of this process. After all, theories or hypotheses about what is efficient are routinely tested by measuring what firms actually do, which indicates a belief in the basic conjecture of the survivor principle. In transaction cost economics, for example, the hypothesis that vertical integration is more efficient than market governance when asset specificity is high is typically tested by measuring whether firms actually integrate when asset specificity is high (David & Han, 2004; Klein, 2005; Shelanski & Klein, 1995), not by examining the performance of alternative integration strategies.<sup>8</sup> In agency theory, hypotheses about the relative efficiency of alternative contracts are tested by measuring which contracts firms actually employ (e.g., Anderson, 1985; Eisenhardt, 1985; Zenger & Marshall, 2000).

If the survivor principle holds, relatedness can be measured by observing which industries firms in competitive markets most often combine. In other words, related industries are defined as those most often performed together. Specifically, we estimate how much the frequencies of actual combinations of 4-digit SIC industries deviate from what one would expect if diversification patterns were random. We take this difference to constitute an SB measure of the relatedness between a pair of industries.<sup>9</sup>

SB measures of relatedness have several potential advantages. First, SB measures incorporate the knowledge of those making portfolio decisions, presumably those with the best knowledge of the relevant benefits and costs of combinations. Even if this knowledge is imperfect, or firms are rife with agency problems, these decisions have been screened by the competitive process, which tends to reverse poor decisions. Second, among the feasible alternatives, the SB approach is holistic and flexible. It is holistic in the sense that it potentially captures all aspects of relatedness that are important for competitive outcomes, including the likelihood that a specific combination meets the conditions of excess capacity, spillovers, and transaction costs noted above. Note here that the screening for these conditions is not explicit, as would be the case ideally. Instead it is implicit, based on the assumption that if these conditions are important for competitive outcomes then decision makers and the competitive process will tend to weed out combinations where the conditions are not satisfied. The approach is flexible in the sense that it allows the causes of relatedness to vary across situations. Whether relatedness is primarily due to production assets, distribution assets, R&D spillovers, or whatever, is left unspecified. It is only assumed that combination frequencies reflect the important sources of relatedness, whatever they may be in each case.

The advantages noted above suggest that an SB approach can measure relatedness effectively, particularly in large-sample empirical research where detailed analysis of strategic decisions is not feasible. But the SB approach has potential disadvantages as well. Managers make mistakes or fail to act in firm owners' interests, and the competitive selection process is not perfect. For this reason, an SB measure will include noise. Critics of the survivor principle believe this noise is substantial (Elster, 1989; Hodgson, 1993; Winter, 1971). All measures of relatedness are noisy, however. The more interesting question is whether SB measures are noisier than feasible alternatives, and even if so, whether SB measures offer other compensating strengths. Empirical analysis is the best judge. Below, we examine this question by comparing SB measures to the most commonly used measure (SIC distances) on the ability to predict the choice of destination industry for firms making diversification decisions and the ability to predict the fate of those diversification decisions. Our tests essentially constitute a horse race in criterion validity between the two approaches. But there is another, perhaps even more profound, problem: The flexibility of the SB approach comes at a cost. In particular, we do not know precisely what the measure captures. The empirical performance of the SB measures can therefore be inflated for reasons unrelated to relatedness. In other words, discriminant validity is a potential problem. Below we discuss the two mechanisms that seem to pose the biggest threat to discriminant validity, namely, herd behavior and mutual forbearance. We think our results suggest that SB measures are sufficiently promising to warrant further work. Moreover, the problem of knowing what exactly a relatedness measure is capturing is shared, at least partially, with SIC-distance measures. The SIC manual itself states: "The classification is organized to reflect the structure of the US economy. It does not follow any single principle, such as end use, nature of raw materials, product or market structure" (SIC Manual, 1987: 699).

Our approach is based on a procedure originally developed by Teece et al. (1994). Let the universe of diversified firms consist of  $K$  firms, each active in two or more of  $I$  industries. Let  $C_{ik} = 1$  if firm  $k$  is active in industry  $i$ . The number of industries participated in by firm  $k$  is  $m_k = \sum_i C_{ik}$ , and the number of diversified firms present in industry  $i$  is  $n_i = \sum_k C_{ik}$ . Let  $J_{ij}$  be the number of diversified firms active in both industries  $i$  and  $j$ , such that  $J_{ij} = \sum_k C_{ik} C_{jk}$ . Thus,  $J_{ij}$  is a count of how often industries  $i$  and  $j$  are actually combined within the same firm.  $J_{ij}$  will be larger if industries  $i$  and  $j$  are related but will also increase with  $n_i$  and  $n_j$ . To remove the effect of the size of industries  $i$  and  $j$ , the number  $J_{ij}$  is compared with the number of expected combinations if diversification patterns were random.

The random diversification hypothesis can be operationalized as a hypergeometric situation where a sample of size  $n_i$  is drawn (without replacement) from a population of  $K$  firms. Those chosen are considered active in industry  $i$ . A second independent sample of size  $n_j$  is then drawn from the population the population of  $K$  firms. Those chosen are considered active in industry  $j$ . The number  $x_{ij}$  of firms active in both  $i$  and  $j$  is then a hypergeometric random variable with population  $K$ , special members  $n_i$ , and sample size  $n_j$ . The distribution function for this variable is then

$$\Pr(X_{ij} = x) = f_{hg}(x, K, n_i, n_j) = \frac{\binom{n_i}{x} \binom{K - n_i}{n_j - x}}{\binom{K}{n_j}}. \quad (1)$$

The mean and variance of  $X_{ij}$  are, respectively,

$$\mu_{ij} = E(X_{ij}) = \frac{n_i n_j}{K}, \quad (2)$$

$$\sigma_{ij}^2 = \mu_{ij} \left(1 - \frac{n_i}{K}\right) \left(\frac{K}{K-1}\right). \quad (3)$$

A standardized measure of the relatedness between industries  $i$  and  $j$  is then constructed based on the difference between  $J_{ij}$  and  $\mu_{ij}$  in the following fashion:

$$SR_{ij} = \frac{J_{ij} - \mu_{ij}}{\sigma_{ij}}. \quad (4)$$

The measure  $SR_{ij}$  is thus a standardized measure of how much the actual number of combinations exceeds expected combinations under the random diversification hypothesis. With this fundamental measure of the relatedness between a pair of businesses, it is possible to compute various relatedness measures.

The two variants chosen here reflect our choice of dependent variables, which are (a) the probability that a given diversified firm will enter a given industry and (b) the probability that a recently entered industry will be exited again. Thus, we wish to develop measures that cast light on the relatedness between a given industry  $i$  (a candidate for entry or a recently entered industry) and the industries present in the parent portfolio.

The first measure captures the weighted average relatedness of the focal industry  $i$  to all other businesses in the parent portfolio. Assume a diversified firm that participates in  $m$  industries. Its business in industry  $j$  has sales of  $s_j$  and SB relatedness  $SR_{ij}$  with industry  $i$ . The weighted average SB relatedness (SURVTOT) of the target industry  $i$  to all other businesses in the firm is then defined as

$$\text{SURVTOT}_i = \frac{\sum SR_{ij} s_j}{\sum s_j}. \quad (5)$$

A parallel measure based on SIC distances can be obtained as follows:

$$\text{SICTOT}_i = \frac{\sum d_{ij} s_j}{\sum s_j}, \quad (6)$$

where

$d_{ij} = 2$  if  $i$  and  $j$  are in the same 3-digit SIC codes,

$d_{ij} = 1$  if  $i$  and  $j$  are in different 3-digit but the same 2-digit SIC codes, and

$d_{ij} = 0$  if  $i$  and  $j$  are in different 2-digit SIC codes.

An alternative approach does not consider how related the focal industry  $i$  is to *all* other business in the corporate portfolio but how related it is to the *two closest* neighboring businesses of the parent. The approach here is to rank the SB measure  $SR_{ij}$  between the focal industry  $i$  and all other industries in the parent portfolio. The two industries with the highest

measure of  $SR_{ij}$  are considered the neighboring businesses. Let  $\lambda_{ij} = 1$  for a business that is defined as a neighbor to business  $i$ , and  $\lambda_{ij} = 0$  for those that are not. The weighted average relatedness of neighbors to business  $i$  (SURVNBOR) is then defined by

$$\text{SURVNBOR}_i = \frac{\sum SR_{ij} s_j \lambda_{ij}}{\sum s_j \lambda_{ij}}. \quad (7)$$

Again, we also computed a parallel measure based on SIC distances:

$$\text{SICNBOR}_i = \frac{\sum d_{ij} s_j \lambda_{ij}}{\sum s_j \lambda_{ij}}, \quad (8)$$

where

- $d_{ij} = 2$  if  $i$  and  $j$  are in the same 3-digit SIC codes,
- $d_{ij} = 1$  if  $i$  and  $j$  are in different 3-digit but the same 2-digit SIC codes,
- $d_{ij} = 0$  if  $i$  and  $j$  are in different 2-digit SIC codes,
- $\lambda_{ij} = 1$  if business  $j$  is defined as a neighbor to business  $i$ ,<sup>10</sup> and
- $\lambda_{ij} = 0$  if business  $j$  is not defined as a neighbor to business  $i$ .

Given these four measures, SURVTOT, SICTOT, SURVNBOR, and SICNBOR, and given the assumption that a superior measure of relatedness will predict diversifying firms' choice of destination industry better, we can formulate the following hypotheses about entry:

*Hypothesis 1:* SURVTOT will explain the choice of destination industry better than SICTOT.

*Hypothesis 2:* SURVNBOR will explain the choice of destination industry better than SICNBOR.

*Hypothesis 3:* SURVTOT and SURVNBOR will both explain the choice of destination industry better than SICTOT and SICNBOR.

## A Test of Survivor-Based Measures of Relatedness

### Method

The test of Hypotheses 1–3 involves two distinct empirical operations. First, we calculate the SB measure of relatedness  $SR_{ij}$  for all possible industry pairs in the U.S. economy. Using this measure, we calculate the SB measures SURVTOT and SURVNBOR for any specific business belonging to any specific parent and for any potential destination industry for any given parent. The second empirical operation is to test our hypotheses comparing the two SB measures and their SIC-based equivalents to the probability of a given parent entering a given industry. We use different samples for each of these operations.

*Calculating  $SR_{ij}$ .* To calculate  $SR_{ij}$ , we use the AGSM/Trinet Large Establishment Database (Trinet). The Trinet database contains biannual records of all U.S. establishments with more than 20 employees, including variables such as 4-digit SIC code, corporate ownership, and sales.<sup>11</sup> Trinet covers the period from 1979 through 1989, but changes in coding practices by Trinet leaves us with usable data from 1981, 1983, 1985, and 1987. By aggregating the establishments for each parent in each 4-digit SIC code, and the different 4-digit

SIC codes for each parent, and different parents for each 4-digit SIC industry, we are able to get a comprehensive picture of diversification patterns in the U.S. economy. Comparison with the Census of Manufacturers indicate that Trinet contains 95% of all establishments it should (Voigt, 1993) and that omissions are most likely for small firms (which are less likely to be diversified).

The primary measure of  $SR_{ij}$  is calculated from the Trinet files of 1981, using all recorded firms active in two or more 4-digit SIC codes as a basis. After deleting single-business firms and government-owned and nonprofit industries, we have 13,164 diversified firms, active in 929 different industries, covering a total of 57,647 individual businesses. Of the 431,056 possible industry pairs, 122,105 are observed. The measure of  $SR_{ij}$  between the observed industry pairs ranges from  $-7.97$  to  $93.55$ , with a mean of  $4.33$  and a standard deviation of  $5.06$ . We also calculate  $SR_{ij}$  for industry pairs that were not combined by 1981, because some of these were combined in the subsequent periods where we observe entry or nonentry and because some of the randomly chosen nonentries create a need to calculate  $SR_{ij}$  for industry pairs that were never combined—for unobserved combinations, we simply set the number of observed combinations ( $J_{ij}$ ) to zero in the formula for calculating  $SR_{ij}$ . Based on these calculations of  $SR_{ij}$ , we calculate measures of SURVTOT and SURVNBOR<sub>*i*</sub> by following the procedures described in the previous section. Finally, note that relatedness between industries as measured by  $SR_{ij}$  changes little over the period covered in this study. The correlation between  $SR_{ij}$  in 1981 and 1983 is  $.941$  and between 1981 and 1985 is  $.895$ .

*Sample for testing Hypotheses 1–3.* The sample for testing Hypotheses 1–3 is derived as follows: To identify a set of firms that entered new industries, we start with the 13,164 diversified firms from the 1981 Trinet file, as described in the preceding paragraph. Next, to obtain the necessary data for the variables of interest, we merge the Trinet data with financial data from the Compustat database. Because the parent identity numbers in these two databases are different, the matching had to be done alphanumerically by parent name. Spelling differences between the two databases resulted in undisputable matches for 1,417 companies in the 1981 files of the two databases. We further require that the sample firms remain present in the Trinet databases throughout 1985. This is because registering entry (or nonentry) requires continued presence in the Trinet database. Because Trinet is biannual, we record entry as having occurred if a firm is active in an industry in the 1983 or 1985 files and that it was not active in the 1981 files. The requirement of continued presence in the Trinet database brings the sample down to 1,097 firms. Lastly, we require that the firms in our sample enter at least one new 4-digit industry in the period 1981 to 1985. This reduces the sample to 594 parent firms. Admittedly, this sampling procedure introduces a survivor bias to our sample. When we still choose to do so, it is because our goal is not to evaluate why or when firms diversify but to compare the ability of different relatedness measures to predict the choice of destination industry by firms actually making a diversification decision.<sup>12</sup>

These 594 firms operated 8,588 businesses in 913 different 4-digit SIC codes in 1981. They entered 2,621 industries between 1981 and 1985, exited 2,117 industries, and remained in 6,404 industries throughout the period.

To test our hypotheses, we included all 2,621 instances of diversifying entry, but rather than using the entire set of potential industries that were not entered by our diversifying

firms, we randomly select a matched sample of these nonentered industries. Our test is designed to compare the relatedness measures on the ability to discriminate between the entered and not-entered industries for a given firm. State-based sampling has been suggested as preferable to a pure random sample when a population is overwhelmingly characterized by one state and will provide unbiased and consistent coefficients for all variables except the constant term (McFadden & Manski, 1981). Adjusting for 15 cases for which data were missing, this resulted in a final sample of a total of 5,227 observations with 2,619 entries into new industries and 2,608 potential entries that did not take place.

*Statistical methods.* To test Hypotheses 1–3, we develop a model of the relationship between the probability of entry and relatedness, which controls for a number of industry and parent variables. The general model is the following:<sup>13</sup>

$$P(\text{entry} = 1) = \beta_1 + \beta_2(\text{industry growth}) + \beta_3(\text{industry concentration}) + \beta_4(\text{industry profitability}) + \beta_5(\text{parent size}) + \beta_6(\text{parent diversity}) + \beta_7(\text{parent profitability}) + \beta_8(\text{parent liquidity}) + \beta_9(\text{parent leverage}) + \beta_{10}(\text{parent relatedness}) + \epsilon.$$

### *Industry-Level Control Variables*

*Industry growth.* High-growth industries should be attractive candidates for entry, as growth allows firms to prosper without having to capture customers from competitors. Thus, industry growth tends to soften competitive rivalry and raise the average profitability (Kwoka & Ravenscraft, 1986; Salinger, 1984; Schmalensee, 1989). High growth may therefore function as a substitute for close relatedness in the eyes of a decision maker who is concerned with postentry performance. Moreover, decision makers may obtain private benefits from growth, which may create a bias toward entering high-growth industries. For these reasons, one would expect industry growth to increase the probability that an industry is chosen as the destination industry. The variable *industry growth* is derived by estimating the growth in percentage of industry sales between 1981 and 1985, as reported in Trinet.

*Industry concentration.* Industry concentration is associated with entry barriers, both structural and strategic (Bain, 1956; Porter, 1980). Entry barriers in turn are by definition negatively related to the probability that an industry is chosen as the target for entry. The variable *industry concentration* is derived by estimating the four-firm concentration ratio of each industry for 1981, based on Trinet data.

*Industry profitability.* The relationship between entry and industry profitability is ambiguous. High profitability invites entry but is more likely when entry barriers are high (Baumol, Panzar, & Willig, 1982); the net effect could go either way. To control for industry profitability, we calculate a measure of the median return on assets (ROA) for each industry over the 1980–1982 period. To calculate profitability, we use all single-business firms in the Compustat industrial file for each 4-digit SIC industry, along with corresponding segments in the Compustat business segment file. Because of incomplete asset allocation, the profitability

ratios in the segment file are systematically higher than the profitability ratios in the corporate database, so we standardize the observations from each database by computing them as percentage deviations from the database mean.<sup>14</sup>

### *Parent-Level Control Variables*

Besides industry-level controls, we also include several parent-level control variables. The reason for doing so warrants some comment. Recall that our main purpose is to see how well different relatedness measures identify destination industries chosen by diversifying firms, not to examine the drivers of diversification decisions. Because we employ a state-based sampling technique in which each industry entered by a given parent is matched with an industry not entered by the same parent, the coefficients on the parent-level variables do not convey information about how they influence the decision to diversify. However, we include parent-level controls to minimize the possibility that parent firm characteristics such as size, diversity, and so on influence the coefficients and relative performance of the relatedness measures we are comparing.

The parent-level controls are *parent size*, measured as total sales; *parent diversity*, measured as the number of SIC codes the parent is active in; *parent profitability*, measured as ROA; *parent liquidity*, measured as the ratio of current assets to current liabilities; and *parent leverage*, measured as long-term debt to market value. All these are computed using 1981 figures.

### *Relatedness Variables*

*Relatedness.* Hypothesis 1 is tested using SURVTOT and SICTOT, which capture the sales-weighted average relatedness of the target industry *i* to *all other* businesses in the parent *k*. We expect the SB version (SURVTOT) to outperform the SIC-based version (SICTOT). (Note that the prefix *SURV* indicates an SB measure, whereas the prefix *SIC* indicates a SIC-based measure.) Hypothesis 2 is tested using SURVNBOR and SICNBOR, which capture the sales-weighted average relatedness between the target industry *i* and the two closest neighboring businesses of the parent. We again expect the SB version (SURVNBOR) to outperform the SIC-based version (SICNBOR). In testing Hypothesis 3, we include all four measures, and we expect the two measures with the prefix *SURV* to outperform both measures with the prefix *SIC*.

*Dependent variable.* Our dependent variable is an indicator variable set equal to 1 if the parent entered a new 4-digit SIC code between 1981 and 1985 and 0 otherwise. The Trinet database is used to identify entries and nonentries.

Variable definitions and data sources are summarized in Table 1. Table 2 shows the means, standard deviations, and correlation coefficients for all independent variables.<sup>15</sup>

**Table 1**  
**Variable Definitions, Data Sources, and Predicted Signs**

Variable	Definition	Data Source
Industry growth	Sales growth in industry <i>i</i> between 1981 and 1985	Trinet
Industry concentration	4-firm concentration ratio in 1981 in industry <i>i</i>	Trinet
Industry profitability	Industry median ROA 1980–1982 in industry <i>i</i>	Compustat
Parent size	Total sales of the parent in 1981	Compustat
Parent diversity	Number of 4-digit SIC codes participated in by parent in 1981	Trinet
Parent profitability	Parent ROA in 1981	Compustat
Parent liquidity	The ratio of current assets to current liabilities in 1981	Compustat
Parent leverage	Long-term debt to market value in 1981	Compustat
SURVTOT	Weighted average <i>survivor-based</i> relatedness of industry <i>i</i> to all industries in the portfolio of the parent	Trinet
SICTOT	Weighted average <i>SIC-based</i> relatedness of industry <i>i</i> to all industries in the portfolio of the parent	Trinet
SURVNBOR	Weighted <i>survivor-based</i> relatedness of industry <i>i</i> to the two closest businesses of the parent	Trinet
SICNBOR	Weighted <i>SIC-based</i> relatedness of industry <i>i</i> to the two closest businesses of the parent	Trinet

Note: ROA = return on assets; SIC = Standard Industrial Classification.

## Results

The results from five logistic regressions are presented in Table 3. Model 1 contains control variables only. Model 2 contains control variables plus the SB measure SURVTOT. Model 3 contains control variables plus the equivalent SIC-based measure SICTOT. Model 4 contains control variables plus the SB measure SURVNBOR. Model 5 contains control variables plus the equivalent SIC-based measure SICNBOR.

As shown in Table 3, all four relatedness measures are positive and statistically significant at the 1% level. This, and the substantial increases in all measures of model performance when any measure of relatedness is included, supports the general hypothesis that relatedness is an important determinant of the direction of diversification.

Hypothesis 1, which predicts that Model 2 will explain the choice of destination industry better than Model 3, is strongly supported; all measures of model performance improve substantially when SURVTOT is substituted for SICTOT. The model  $\chi^2$  improves from 677.81 to 1,808.84, an improvement that is statistically significant at the 1% level. The pseudo  $R^2$  measure (Nagelkerke) increases from 16.8% to 40.4%. These increases represent improvements of approximately 167% and 140%, respectively. The ability to predict entries and nonentries also increases from 63.0% to 75.9%, an improvement of 99% (over the 50% base rate). We interpret these results as strong support for Hypothesis 1.

Hypothesis 2 predicts that Model 4 explains the choice of destination industry better than Model 5. This hypothesis is also strongly supported. Again, all measures of model performance improve substantially when the measure SURVNBOR is substituted for SICNBOR.

**Table 2**  
**Means, Standard Deviations, and Correlation Coefficients of Independent Variables**

Variables	Mean	Standard Deviation	Industry Growth	Industry Concentration	Industry Profitability	Parent Size	Parent Diversity	Parent Profitability	Parent Liquidity	Parent Leverage	SURVTOT	SICTOT	SURVNBOR	SICTOT SURVNBOR
<b>Industry variables</b>														
Industry growth	0.58	1.52												
Industry concentration	31.68	21.63	.16***											
Industry profitability	0.11	0.47	.01	-.03**										
<b>Firm variables</b>														
Parent size	24.221	44.475	.01	.04***	-.04***									
Parent diversity	25.49	21.79	-.00	.04***	-.05***	.55***								
Parent profitability	0.067	0.045	.01	.02	.03**	.04***	.03**							
Parent liquidity	2.08	0.85	.02	-.01	.03**	-.31***	-.27***	.23***						
Parent leverage	0.18	0.11	-.02	-.03**	-.02	-.10***	-.02	-.41***	-.11***					
<b>Relatedness variables</b>														
SURVTOT	4.22	6.07	.01	-.09***	.01	-.10***	-.15***	.00	.08***	-.09***				
SICTOT	0.10	0.27	.01	.00	.02	-.08***	-.12***	-.02	.08***	-.08***	.59***			
SURVNBOR	11.18	9.60	.01	-.10***	.00	.10***	.20***	.02	-.06***	-.06***	.73***	.40***		
SICNBOR	0.51	0.68	.01	-.04***	.02*	.12***	.26***	.02	-.05***	-.03*	.38***	.54***	.55***	

Note: N = 5,227. See Table 1 for definitions of the variables. Statistically significant correlations: \* = 10% level; \*\* = 5% level; \*\*\* = 1% level.

**Table 3**  
**Relatedness and the Probability of Entry**

	1	2	3	4	5
Constant	0.522*** (0.136)	-0.792*** (0.169)	0.310** (0.143)	-0.935*** (0.163)	0.255** (0.145)
Industry growth	0.060*** (0.020)	0.060*** (0.021)	0.066*** (0.020)	0.045* (0.023)	0.058*** (0.020)
Industry concentration	-0.018*** (0.001)	-0.016*** (0.002)	-0.020*** (0.002)	-0.016*** (0.002)	-0.019*** (0.002)
Industry profitability	0.044 (0.062)	-0.067 (0.073)	-0.032 (0.065)	-0.029 (0.073)	-0.047 (0.067)
Parent size	0.000 (0.000)	0.000* (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Parent diversity	0.001 (0.002)	0.005*** (0.002)	0.002 (0.002)	-0.017*** (0.002)	-0.012*** (0.002)
Parent profitability	0.020 (0.747)	-0.029 (0.941)	0.162 (0.778)	-0.507 (0.900)	-0.066 (0.791)
Parent liquidity	0.001 (0.037)	-0.113** (0.047)	-0.050 (0.040)	-0.004 (0.043)	-0.025 (0.039)
Parent leverage	-0.094 (0.304)	0.796** (0.369)	0.136 (0.318)	0.291 (0.358)	-0.055 (0.325)
SURVTOT		0.368*** (0.012)			
SICTOT			4.116*** (0.256)		
SURVBOR				0.179*** (0.006)	
SICNBOR					1.433*** (0.056)
-2 log likelihood	6,807.5	5,139.0	6,307.7	5,218.95	5,962.51
Model $\chi^2$	178.05***	1,808.84***	677.81***	1,766.57***	1,023.01***
$\Delta\chi^2$ vs. Model 1		1,632.93***	499.77***	1,588.52***	844.96***
$\Delta\chi^2$ Model 2 vs. Model 3		1,133.16***			
$\Delta\chi^2$ Model 4 vs. Model 5				743.56***	
Nagelkerke $R^2$	0.046	0.404	0.168	0.394	0.245
Correct classifications	56.9%	75.9%	63.0%	76.1%	68.6%

Note: Logistic regressions indicate the probability that a sample firm will enter a new industry. Standard errors are in parentheses. See Table 1 for definitions of the variables.  $N = 5,227$ . Statistical significance: \* = 10% level; \*\* = 5% level; \*\*\* = 1% level.

The model  $\chi^2$  improves from 1,023.01 to 1,766.57, an increase that is statistically significant at the 1% level. The pseudo  $R^2$  measure increases from 24.5% to 39.4% (Nagelkerke). These represent improvements of about of 73% and 61%, respectively. The ability to predict entries and nonentries also increases from 68.6% to 76.1%, an improvement of 40% (over the 50% base rate). We interpret these findings as strong support for Hypothesis 2.

Hypothesis 3 predicts that both SB measures will outperform both SIC-based measures. This hypothesis is supported if the lowest performing SB measure outperforms the best performing SIC-based measure. As shown in Table 3, the lowest performing SB measure is SURVBOR, whereas the best performing SIC-based measure is SICNBOR. The test of Hypothesis 3 therefore reduces to the same test as Hypothesis 2, which as noted was strongly supported. We therefore conclude that Hypothesis 3 is also strongly supported.

We also examined possible complementarity between the SIC-based and SB measures by checking to see if adding SIC measures to a model including SB measures increases model performance. In these cases, the SIC measures are statistically significant but not economically

significant. For example, adding SICTOT to a model with SURVTOT increases Nagelkerke pseudo  $R^2$  from 40.4% to 40.7%, and adding SICNBOR to a model with SURVNBOR increases Nagelkerke pseudo  $R^2$  from 39.4% to 40.4%.

### An Additional Test

The empirical analysis in the previous section provides strong evidence that an SB measure of relatedness captures whatever it is that decision makers are acting on better than an SIC-based measure does. However, entry decisions are about expected performance, whereas the relatedness construct is meant to capture actual efficiency consequences. Expected performance and actual efficiency are not necessarily the same. To see if SB is linked closely to actual efficiency, we examine the fate of diversifying firms' entry decisions. Our assumption is that the ability to survive in a newly entered industry is closely related to efficiency. In other words, inefficient combinations are systematically more likely to result in exit than are efficient ones.

If the SB measure is related to efficiency, one would expect it to be negatively related to the probability of exit, and furthermore, if it is a superior measure of relatedness, one would expect it to explain the probability of exit better than SIC-based measures do. This suggests the following hypothesis:

*Hypothesis 4:* SB measures of relatedness are negatively related to the probability of exit and are better than SIC-based measures in explaining the probability of exit.

#### *The Empirical Test*

To test Hypothesis 4, we do not restrict the sample to unambiguous matches between Trinet and Compustat, because 1987 is the last year for which we have Trinet data on industry participation, which are needed to determine if a firm is still active in an entered industry. Because the competitive process may take time to filter entry decisions, we restrict the sample to entries made between 1981 and 1983. Using data from 1987 to measure postentry survival poses a further challenge: The SIC system was revised in 1987, and there is no unambiguous procedure for converting between the two versions of the SIC system. Fortunately, 609 of the 913 relevant 4-digit SIC codes<sup>16</sup> were unchanged by the revision, so we restrict the sample to entries into these 609 industries. (Requiring an unambiguous match between Compustat and Trinet would have reduced this sample to 215 firms, with a severe bias toward larger firms.) To keep the sample representative, we use all entries recorded in the Trinet database between 1981 and 1983, for which their continued existence in 1987 could be determined. This provides a sample of 2,439 entries by 1,250 parent firms, among which 1,004 were subsequently exited, and 1,435 were not.

We again use a logistic regression to test Hypothesis 4. The dependent variable is the probability that an industry entered between 1981 and 1983 is exited by 1987. The general model is the following:  $P(\text{exit} = 1) = \beta_1 + \beta_2(\text{relatedness}) + \epsilon$ . The independent variables are the four relatedness measures used to test Hypotheses 1–3. Table 4 shows the means, standard deviations, and correlation coefficients of the four relatedness variables.

**Table 4**  
**Means, Standard Deviations, and Correlation**  
**Coefficients of Independent Variables, Exit Sample**

Variable	Mean	Standard Deviation	SURVTOT	SURVNBOR	SICTOT
SURVTOT	8.90	9.99			
SURVNBOR	15.74	11.80	.75***		
SICTOT	0.122	0.26	.46***	.32***	
SICNBOR	0.72	0.72	.18***	.37***	.60***

*Note:*  $N = 5,227$ . See Table 1 for definitions of the variables. Statistically significant correlations: \* = 10% level; \*\* = 5% level; \*\*\* = 1% level.

Because we are simply ranking the ability of the different relatedness measures to explain the probability that recently entered businesses will be exited, rather than studying the coefficient values per se, we do not include control variables in these regressions. The only exogenous variable we expect to affect the performance of the relatedness variables differently in this context is industry concentration, which we explore in detail below. However, as a robustness check, we repeat the analysis reported below using controls for industry growth, industry concentration, industry profitability, parent size, and parent diversity; our findings remain the same.<sup>17</sup>

## Results

Table 5 presents the results of four logistic regression analyses, one for each relatedness measure. The first claim in Hypothesis 4 is that the SB measures are negatively related to the probability of exit. As can be seen from the first row of Table 5, both SB measures are indeed negatively signed and significant at the 1% level. The second claim was that the SB measures would explain the probability of exit better than their SIC-based counterparts do. As Table 5 shows, this claim is supported. The  $\chi^2$  and the pseudo  $R^2$  are nearly three times as large in the model including SURVTOT compared to the model with SICTOT. Furthermore, the model including SURVNBOR outperforms the model with SICTOT. The difference in  $\chi^2$  between the models containing these two measures is statistically significant at the 1% level. The ability of the measure SICNBOR to predict exit is statistically significant but seems economically insignificant.

Again, we ask whether the measures appear to capture different, complementary aspects of relatedness. To explore this, we run a model with both SURVTOT and SICTOT included. The results are presented in column 5 of Table 5. As Table 5 shows, SICTOT adds little to a model containing SURVTOT only. The increase in model  $\chi^2$  is not significant by conventional standards, nor is the coefficient on SICTOT. We also note that the increase in pseudo  $R^2$  is negligible. In other words, we find little evidence of complementarity between the measures.

In short, the superiority of the SB measures seems to hold for both entry decisions and entrant survival. In fact, although relatedness generally contributes more toward explaining

**Table 5**  
**Relatedness and the Probability of Exit**

	1	2	3	4	5
SURVTOT	-0.052*** (0.005)				-0.048*** (0.006)
SURVNBOR		-0.027*** (0.004)			
SICTOT			-0.641*** (0.106)		-0.211* (0.120)
SICNBOR				-0.154*** (0.058)	
-2 log likelihood	3,184.8	3,249.5	3,264.6	3,297.4	3,181.67
Model $\chi^2$	119.79***	55.07***	39.92***	7.21***	122.93***
$\Delta\chi^2$ vs. SURVTOT only					3.15*
Nagelkerke $R^2$	0.065	0.030	0.022	0.004	0.066

*Note:* Logistic regressions indicate the probability that a sample firm will exit a previously entered industry. Standard errors are in parentheses. See Table 1 for definitions of the variables.  $N = 2,439$ . Statistical significance: \* = 10% level; \*\* = 5% level; \*\*\* = 1% level.

entry decisions than explaining entrant survival, the relative superiority of the SB approach is actually higher for entrant survival than for entry.

## Alternative Interpretations

As noted earlier, a potent argument against employing SB measures of relatedness is discriminant validity. Although the flexibility of the SB approach allows it to capture a variety of drivers of relatedness, we do not know precisely what we are measuring, and such a measure could be contaminated by mechanisms unrelated to relatedness. Our approach to this challenge is to examine two mechanisms that theoretically seem to constitute the biggest threat to discriminant validity: herding and mutual forbearance. We conclude that these two are unlikely to alter the conclusion that the SB approach has substantial net benefits over the SIC-based approach, although we acknowledge that more work is needed to verify that the SB measures capture what we want them to.

### *Herd Behavior*

The entry patterns we observe could reflect herd behavior rather than efficiency. "Rational" herding occurs when decision makers suppress their private information, either because making a bad decision is less costly when others make the same decision (Scharfstein & Stein, 1990) or because decision makers believe that the decisions of others reflect valuable private information (Banerjee, 1992; Bikhchandani et al., 1992). Either way, entry decisions may be based on the actions of others rather than on superior private knowledge about which industries are related to each other. Our SB measures of relatedness are particularly susceptible to contamination from herd behavior, as they measure what combinations firms in the same industries as the focal firm have chosen previously. SIC-based

approaches, on the other hand, are based on standardized distances in the SIC system and are little influenced by the behavior of other firms.<sup>18</sup> Support for Hypotheses 1–3 may therefore be driven by herd behavior rather than by a superior ability to capture relatedness.

However, a herd behavior interpretation is inconsistent with our findings on postentry survival. If the SB measure primarily captures herding, rather than efficiency, we would expect it to predict entry well but survival poorly. Once entry has occurred, competitive forces should begin to screen the good decisions from the bad. As we showed previously, however, relatedness is a good predictor of survival—related industries are less likely to be exited than unrelated ones—and the SB measure predicts survival much better than its SIC-based equivalent does. A possible counterargument is that exit decisions also reflect herding, contaminating our measure of postentry performance. We think this concern is misplaced, however. It implies that those firms least sensitive to herding before entry (those that chose rare combinations) become the most sensitive to herding after entry. The reversal of a previous entry decision seems more likely to result from poor performance than from a sudden change of independence. In other words, the advantages of the SB measures, relative to their SIC-based equivalents, are enough to overcome possible contamination by herding.

### *Mutual Forbearance*

Another potential concern is mutual forbearance. The forbearance hypothesis suggests that contact between firms across markets induces implicit or explicit agreements to refrain from aggressive competition (Edwards, 1955). Multimarket contact enables a firm to respond to an aggressive action by a multipoint rival in one market with a retaliatory action in another. Such behavior increases the potential costs of aggressive moves, decreasing market competitiveness (Karnani & Wernerfelt, 1985).

Firms may therefore prefer to enter industries where they will meet existing competitors as a mean of establishing mutual forbearance (Hypotheses 1–3). Or they may refrain from exiting a weak position in one industry (Hypothesis 4), not because they are reaping gains from relatedness but because they must remain in that industry to dissuade rivals from acting aggressively in other industries. There is some empirical support for the claim that the creation and exploitation of mutual forbearance affects the behavior and patterns of diversification (Greve & Baum, 2001).

SB measures of relatedness are particularly prone to incorporate mutual forbearance motives because they are built by counting the frequencies of multimarket contact across industries. If our analysis is driven by mutual forbearance rather than relatedness, we expect SB measures to work best for predicting entry into more highly concentrated industries but be no better than SIC-based measures for predicting entry into less concentrated industries. The reason is that the mutual forbearance motive requires a minimum level of concentration to be plausible. In fragmented markets, existing or potential competitors without other markets to protect can force multimarket firms into vigorous competition. To test whether mutual forbearance can account for the better performance of SB measures of relatedness, we formulate the following hypothesis:

*Hypothesis 5:* SB measures of relatedness explain the probability of entry better than SIC-based measures do when industry concentration in the destination industry is high but not when it is low.

### *A Final Test*

To test Hypothesis 5, we use the same sample used to test Hypotheses 1–3. We divide the sample into two equally sized subsamples, one containing the most concentrated industries and the other containing the least concentrated industries, and rerun the logistic regressions used to test Hypotheses 1–3 on each sample.

Table 6 presents results for both the high- and low-concentration subsamples. The results strongly contradict Hypothesis 5. Not only are the SB measures superior to the SIC-based measures in both subsamples, but SB measures' relative superiority is highest in the low-concentration subsamples. This is seen most clearly in the difference in  $\Delta\chi^2$  from Model 2 to Model 3 and from Model 4 to Model 5. These changes in  $\Delta\chi^2$  are not only positive and highly significant for both subsamples but are actually larger for the low-concentration subsample than for the high-concentration subsample. In other words, the advantage of substituting SB measures for SIC-based measures is greatest where the mutual forbearance interpretation is least plausible. This suggests that the superior performance of the SB measures of relatedness is unlikely to be caused by mutual forbearance motives.

## Conclusion

Improving our measures of key theoretical constructs is essential for advances in strategic management research. Few constructs are more important than relatedness, and few measurement procedures have been more heavily criticized than those used to capture relatedness (Fan & Lang, 2000; Markides & Williamson, 1994, 1996; Robins & Wiersma, 1995, 2003; Silverman, 1999). The SB approach investigated here allows the actions of firms in competitive markets to tell us which industries are related to which, instead of imposing some a priori view of relatedness on the data. Rather than letting the SIC system or the researcher be the judge of what is related to what, we rely on the wisdom of local decision makers and the screening function of the competitive process.

Of course, this procedure also means that we do not directly know the causes of relatedness in each particular instance. The flexibility of this approach, namely its ability to capture a variety of underlying causes of relatedness between industry combinations, makes it a potentially noisy measure. An important task for subsequent work is to treat SB relatedness as a dependent variable, examining its cross-sectional determinants and its evolution over time. In this way, we may learn better whether the measure captures what it should, and if it does, what relatedness really "is." We can also use SB measures to understand what relatedness "does" by looking at its effects on performance, growth, entry mode, financing, organizational choice, merger-and-acquisition performance, and the like. For example, in a related project, we find a significant positive relationship between relatedness and acquirer performance when relatedness is measured using the SB approach but no relationship when SIC-distance measures were used (Lien & Klein, 2006).

Table 6  
 Relatedness and Probability of Entry, Split Concentration Samples

	1		2		3		4		5	
	High	Low								
Constant	1.178*** (0.223)	0.214 (0.203)	-0.380 (0.276)	-0.827*** (0.248)	1.065*** (0.239)	0.014 (0.212)	-0.486* (0.267)	-1.043*** (0.241)	0.956*** (0.244)	-0.036 (0.214)
Industry growth	0.044* (0.023)	0.118*** (0.040)	0.044 (0.027)	0.120*** (0.042)	0.050*** (0.024)	0.130*** (0.039)	0.035 (0.029)	0.080* (0.044)	0.045* (0.025)	0.113*** (0.041)
Industry concentration	-0.024*** (0.002)	-0.019*** (0.007)	-0.019*** (0.003)	-0.028*** (0.008)	-0.027*** (0.003)	-0.024*** (0.007)	-0.018*** (0.003)	-0.029*** (0.008)	-0.025*** (0.003)	-0.022*** (0.007)
Industry profitability	-0.127 (0.084)	0.275*** (0.094)	-0.183* (0.100)	0.086 (0.110)	-0.240*** (0.090)	0.230*** (0.097)	-0.092 (0.098)	-0.050 (0.110)	-0.250*** (0.093)	0.201* (0.099)
Parent size	0.000 (0.000)									
Parent diversity	0.001 (0.002)	0.001 (0.002)	0.006*** (0.002)	0.004*** (0.003)	0.003 (0.002)	0.001 (0.002)	-0.016*** (0.003)	-0.018*** (0.003)	-0.011*** (0.002)	-0.013*** (0.003)
Parent profitability	-1.178 (1.197)	-0.613 (0.966)	-0.712 (1.474)	0.293 (1.250)	-1.193 (1.282)	0.819 (1.005)	-1.626 (1.433)	-0.139 (1.181)	-1.658 (1.308)	0.720 (1.022)
Parent liquidity	-0.051 (0.053)	0.050 (0.053)	-0.146** (0.069)	-0.085 (0.065)	-0.123** (0.060)	0.014 (0.055)	-0.050 (0.062)	0.035 (0.059)	-0.097* (0.059)	0.038 (0.055)
Parent leverage	-0.624 (0.450)	0.330 (0.418)	0.454 (0.540)	1.032*** (0.510)	-0.413 (0.478)	0.546 (0.433)	-0.106 (0.528)	0.579 (0.494)	-0.711 (0.492)	0.429 (0.441)
SURVTOT			0.357*** (0.017)	0.379*** (0.018)						

(continued)

**Table 6 (continued)**

	1		2		3		4		5	
	High	Low	High	Low	High	Low	High	Low	High	Low
SICTOT					4.248*** (0.348)	4.084*** (0.384)				
SURVNBOR					0.168*** (0.008)	0.189*** (0.009)				
SICNBOR					1.504*** (0.078)	1.373*** (0.081)				
-2 log likelihood	3355.04	3420.41	2543.0	2582.6	3060.13	3208.64	2612.45	2594.89	2873.09	3054.19
Model $\chi^2$	120.06***	25.60***	910.78***	848.68***	414.97***	237.38***	862.65***	851.13***	602.07***	391.83***
Nagelkerke $R^2$	0.0462	0.014	0.407	0.386	0.203	0.121	0.387	0.385	0.284	0.194
$\Delta\chi^2$ vs. controls only			790.57***	822.80***	294.91***	211.78***	742.59***	825.53***	481.94***	366.23***
$\Delta\chi^2$ Model 2 -			High-concentration sample: 495.66***							
$\Delta\chi^2$ Model 3			Low-concentration sample: 611.02***							
$\Delta\chi^2$ Model 4 -									High-concentration sample: 260.65***	
$\Delta\chi^2$ Model 5									Low-concentration sample: 459.30***	

Note: Logistic regressions indicate the probability that a sample firm will enter a new industry. Standard errors are in parentheses. See Table 1 for definitions of the variables.  $N = 2,614$  for the high-concentration sample, and  $N = 2,613$  for the low-concentration sample. Statistical significance: \* = 10% level; \*\* = 5% level; \*\*\* = 1% level.

We have focused here on marginal diversification moves (i.e., entry and exit) rather than on the characteristics of entire portfolios as is more common in the literature. This reflects our belief that the literature on relatedness and diversification may have moved too quickly to focus on aggregate constructs without sufficient understanding of the basic building blocks. The basic element in relatedness is the relationship between industry pairs, and the basic diversification action in corporate strategy is the addition or subtraction of a particular line of business. Of course, portfolio composition involves additional issues, but we think diversification itself is better understood when built on strong microfoundations. Consider, for example, Palich et al.'s (2000) finding of a curvilinear relationship between relatedness and performance.

One possible explanation for this pattern is that some firms run out of related target industries but continue to diversify. Whether firms tend to add unrelated industries only when related target industries are exhausted can be examined using the SB measures discussed here. Another possible source of curvilinearity is that increasing managerial complexity ultimately takes its toll such that at a certain level of diversity the benefits of relatedness decrease. Put differently, one should observe that diversity moderates the benefits of relatedness. Such moderation effects are more easily examined with a continuous, quantitative measure of the fundamental relatedness between industry pairs.

Furthermore, the SB approach is useful for characterizing interindustry relationships more generally. If we think of relatedness as distance, the SB approach allows us to place any industry within the space of all other industries. To see why this is useful, consider Santalo and Becerra's (2008) finding that the diversification–performance link is heterogeneous across industries. Some industries are dominated by diversified firms, whereas others are dominated by specialists. Santalo and Becerra show that the diversification discount arises primarily from diversification into the latter type of industries. In the spatial metaphor of relatedness, specialist industries would be described as those that are distant from all other industries. The SB approach allows us to quantify these kinds of relationships. How closely related is an industry's nearest "neighbor" industry? How many industries are combined with the focal industry? What is the sum of the relatedness to all these industries? Does it matter whether the closest neighboring industry is fragmented or concentrated? And so on.

Finally, note that it is straightforward to compute average relatedness scores for entire portfolios as well as the "portfolio-target industry" measures we have used here. One can also quantify other properties, such as the degree to which the portfolio is linked or clustered, and so on. Indeed, we expect that SB measures of firm-level relatedness may help resolve some difficult issues in the diversification discount literature. As noted above, that literature has made considerable progress in addressing the endogeneity and unobserved heterogeneity problems that plague research on the performance effects of organizational form and strategy. Still, that literature has tended to rely on coarse, distance-based measures of relatedness (or, in the case of the corporate finance literature, ignored relatedness altogether). Introducing a robust, continuous measure like the SB measure allows researchers to deal more easily with nonlinear relationships between diversification and relatedness and performance and to interact relatedness with other characteristics and behaviors.

## Notes

1. In this we include as a special case the situation where the resource in question is a public input, so that excess capacity will always exist.

2. Recent literature emphasizes *dynamic* complementarities, the ability to identify new ways of combining existing resources or speed up the development of new resources. The benefits to similarity in this context arise because such dynamic complementarities may be greater if the industries in question share some basic features (March, 1991) or because some common characteristics facilitate their exploitation (Finkelstein & Halebian, 2002; Prahalad & Bettis, 1986). The degree of dynamic complementarity between industries thus depends on the balance between variety and similarity (Christensen & Foss, 1997). Industries with appropriate balances between variety and similarity produce larger dynamic complementarities than do industries that are too different or too similar. Empirically, this implies that portfolios of businesses with strong interindustry complementarities should be considered related (or “coherent,” in Teece, Rumelt, Dosi, and Winter’s [1994] language) and that firms with related combinations of industries should outperform firms with unrelated combinations and single-business firms, *ceteris paribus* (again, assuming positive contracting costs).

3. Complementarities, combined with indivisibility, also play an important role in Penrose’s (1959) approach to growth as well as in Lachmann’s (1956) theory of capital heterogeneity. Notes Lachmann (1956),

As capital becomes more plentiful its accumulation does not take the form of multiplication of existing items, but that of a change in the composition of capital combinations. . . . The capital structure will thus change . . . almost certainly towards a higher degree of complexity, i.e. more types of capital items will now be included in the combinations. The new items, which either did not exist or were not used before, will mostly be of an indivisible character. *Complementarity plus indivisibility* are the essence of the matter. (pp. 79-80)

4. A study that does consider excess capacity is that by Chatterjee and Wernerfelt (1991).

5. Silverman (1999) explicitly examines the influence of market failure on patterns of diversification.

6. For the view that the competitive process creates outcomes that are optimizing, see Friedman (1953).

7. Friedman (1953) is a possible exception.

8. Some newer articles in the empirical transaction cost literature do not assume the survivor principle but, rather, employ a two-stage procedure in which the relationship between transactional characteristics and governance structure is endogenously chosen in the first stage and then used to explain performance in the second stage. Silverman, Nickerson, and Freeman (1997), for example, show that transaction cost efficiency is positively correlated with firm survival in the for-hire trucking industry, whereas Bigelow (2001) examines outsourcing arrangements in the U.S. automobile industry and finds that transactions that are appropriately aligned tend to last longer than inappropriately organized ones.

9. The survivor principle has hardly been immune from criticism, as used both for theory building and theory testing. Our use of the survivor principle assumes that the competitive selection process is effective at screening for resource combinations that generate efficiencies. Critics have noted that Friedman’s optimizing version of the survivor principle suffers from the problem of sufficient variation: The selection process can only choose the optimal behavior or decision if that behavior is part of the set of initial behaviors (Nelson & Winter, 1982), unless there is entrepreneurial experimentation and learning. Moreover, unless incremental change is continually beneficial—that is, there is no local optimum next to the global optimum—it may be impossible to reach the optimal solution through small evolutionary steps (Elster, 1989). Alchian’s comparative-efficiency version of the survivor principle—a survival of the fitter, not the fittest—is not vulnerable to this critique. However, this more modest version is not immune to some of the other criticisms of the survivor principle. Winter (1964, 1971), for instance, points out that because of environmental change, selection has a moving target; if environmental conditions change faster than the operation of the selection and adaptation process, it is difficult to say which environmental conditions a population is adapted to. A related objection is that the competitive process selects for overall firm performance, not the individual decisions that determine performance (Elster, 1989).

10. We resorted to the following procedure when businesses were equidistant from the target industry: First, we identified the closest neighbor; if several were equidistant, we used the sales of the *largest* business to compute sales weights. Then we identified the second closest. If there were several second-closest firms, we chose the *smallest* of

these. This was done to reflect our assumption that one would expect closer cooperation with the closest neighbor than with the second closest. The procedure implies that when the closeness to the two neighbors differed, we weighted our measure in favor of the closest of the two.

11. The Trinet database contains an unknown proportion of sales figures that are imputed from multiplying employee counts with average industry sales per employee. To examine whether this constitutes a substantial source of error for our sample, we correlated the sales data from Trinet with Compustat data. This resulted in a correlation of .893, which indicates that the sales data in Trinet are of acceptable quality.

12. Note that for all findings our hypotheses are unchanged if we instead use the full Trinet database (i.e., we do not match with Compustat), or if we do not restrict the sample to surviving firms, or if we remove the criterion of having entered at least one new industry. However, we do note that the latter criterion does increase the mean size of the firms in the sample significantly, from 69.1 in the Compustat population ( $SD = 3,315$ ) to 206.5 in our sample ( $SD = 6,955$ ).

13. We also experimented with including measures of liquidity and of leverage at the parent level. Non of these were significant, which is not surprising. One would assume that these variables affect the decision to diversify or not and, to a much lesser degree, the decision about which industry to target given that one is making a diversification decision.

14. Following Berger and Ofek (1995), we define industries at the 4-digit Standard Industrial Classification (SIC) level where five or more observations are available at that level, at the 3-digit SIC level if five or more observations are available only at that level, and at the 2-digit SIC level if five or more observations are available only at that level. For a critique of this approach, see Santalo and Becerra (2008).

15. One possibly surprising finding from Table 2 is that the two survivor-based (SB) measures are more closely correlated with each other than SURVTOT is correlated with SICTOT and SURVNBOR is correlated with SICNBOR. A contributing factor is the fact that for firms in a small number of industries, the correlation between the two SB measures approaches 1 by construction. For firms with 2 businesses, the correlation is 1.0; for firms with 3 businesses, it is .96. But even disregarding firms with small portfolios, the correlation between the two SB measures is high; if we restrict the sample to cases involving firms with more than 10 businesses, the correlation is .7. The reason for this high correlation is the widespread tendency to construct related portfolios. If a potential target industry is close to some industries in a firm's portfolio—that is, if the two closest industries are close—then this target industry is systematically more likely to have a low average distance to all the industries in the portfolio. Conversely, if the target industry is distant from the two closest neighboring industries, then it is systematically likely to have a high average distance to all industries in the portfolio.

A useful analogy here is that diversification patterns create industries that clustered like stars in galaxies. A star is likely to have much lower minimum and average distance to other stars in the same galaxy than to stars in other galaxies.

16. SIC codes referring to public and nonprofit industries are omitted.

17. Regressions with control variables included are available from the authors.

18. Many industries that are close in the SIC system are never actually combined, for example, whereas many seemingly distant industries are frequently combined.

## References

- Alchian, A. A. 1950. Uncertainty, evolution, and economic theory. *Journal of Political Economy*, 58: 211-221.
- Alchian, A. A. 1969. Corporate management and property rights. In H. Manne (Ed.), *Economic policy and the regulation of corporate securities*: 337-360. Washington, DC: American Enterprise Institute.
- Amihud, Y., & Lev, B. 1981. Risk reduction as a managerial motive for conglomerate merger. *Bell Journal of Economics*, 12: 605-617.
- Anderson E. 1985. The salesperson as outside agent or employee: A transaction cost analysis. *Marketing Science*, 4: 234-254.
- Bain, J. S. 1956. *Barriers to new competition, their character and consequences in manufacturing industries*. Cambridge, MA: Harvard University Press.
- Banerjee, A. V. 1992. A simple model of herd behavior. *Quarterly Journal of Economics*, 107: 797-817.

- Baumol, W. J., Panzar, J. C., & Willig, R. D. 1982. *Contestable markets and the theory of industry structure*. New York: Harcourt.
- Berger, P. G., & Ofek, E. 1995. Diversifications effect on firm value. *Journal of Financial Economics*, 37(1): 39-65.
- Bigelow, L. 2001. *Efficient alignment and survival in the U.S. automobile industry*. Working paper, Olin School of Business, Washington University, St. Louis, MO.
- Bikhchandani, S., Hirshleifer, D., & Welch, I. 1992. A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy*, 100: 992-1026.
- Breschi, S., Lissoni, F., & Malerba, F. 2003. Knowledge-relatedness in firm technological diversification. *Research Policy*, 32(1): 69-87.
- Campa, J. M., & Kedia, S. 2002. Explaining the diversification discount. *Journal of Finance*, 57: 1731-1762.
- Caves, R. E., Porter, M. E., & Spence, A. M. 1980. *Competition in the open economy: A model applied to Canada*. Cambridge, MA: Harvard University Press.
- Chatterjee, S. 1990. Excess resources, utilization costs, and mode of entry. *Academy of Management Journal*, 33: 780-800.
- Chatterjee, S., & Blocher, J. D. 1992. Measurement of firm diversification: Is it robust? *Academy of Management Journal*, 35: 874-888.
- Chatterjee, S., & Wernerfelt, B. 1991. The link between resources and type of diversification: Theory and evidence. *Strategic Management Journal*, 12: 33-48.
- Chevalier, J. A. 2000. *Why do firms undertake diversifying mergers? An examination of the investment policies of merging firms*. Working paper, Graduate School of Business, University of Chicago, IL.
- Christensen, J. F., & Foss, N. J. 1997. Dynamic corporate coherence and competence based competition: Theoretical foundations and strategic implications. In A. Heene & R. Sanchez (Eds.), *Competence-based strategic management*. Chichester, UK: Wiley.
- David, R. J., & Han, S.-K. 2004. A systematic assessment of the empirical support for transaction cost economics. *Strategic Management Journal* 25: 39-58.
- Edwards, C. D. 1955. Conglomerate bigness as a source of power. In *NBER conference report on business concentration and economic policy*. Princeton, NJ: Princeton University Press.
- Eisenhardt, K. M. 1985. Control: Organizational and economic approaches. *Management Science* 31: 134-149.
- Elster, J. 1989. *Nuts and bolts for the social sciences*. New York: Cambridge University Press.
- Engelsman, E. C., & van Raan, A. F. J. 1992. A patent-based cartography of technology. *Research Policy*, 23: 1-26.
- Fan, J. P. H., & Lang, L. H. P. 2000. The measurement of relatedness: An application to corporate diversification. *Journal of Business*, 73: 629-660.
- Farjoun, M. 1994. Beyond industry boundaries: Human expertise, diversification and resource-related industry groups. *Organization Science*, 5: 185-199.
- Finkelstein, S., & Halebian, J. 2002. Understanding acquisition performance: The role of transfer effects. *Organization Science*, 13: 36-47.
- Foss, N. J., & Christensen, J. F. 2001. A market-process approach to corporate coherence. *Managerial and Decision Economics*, 22: 213-226.
- Friedman, M. 1953. The methodology of positive economics. In *Essays in positive economics*. Chicago: University of Chicago Press.
- Gertner, R. H., Scharfstein, D. S., & Stein, J. C. 1994. Internal versus external capital markets. *Quarterly Journal of Economics*, 109: 1211-1230.
- Gimeno J. 1999. Reciprocal threats in multimarket rivalry: Staking out "spheres of influence" in the US airline industry. *Strategic Management Journal*, 20: 101-128.
- Graham, J., Lemmon, R. M., & Wolf, J. 2002. Does corporate diversification destroy value? *Journal of Finance*, 57: 695-720.
- Greve, H. R., & Baum, J. A. C. 2001. Introduction: A multiunit, multimarket world, *Advances in Strategic Management*, 18: 1-28.
- Hall, E. H., & St. John, C. H. 1994. A methodological note on diversity measurement. *Strategic Management Journal*, 15: 153-168.
- Hill, C. W. L., & Hoskisson, R. E. 1987. Strategy and structure in the multiproduct firm. *Academy of Management Review*, 12: 331-341

- Hodgson, G. 1993. *Economics and evolution*. Ann Arbor, MI: University of Chicago Press.
- Hoskisson, R. E. & Hitt, M. A. (1990). Antecedents and performance outcomes of diversification: A review and critique of theoretical perspectives. *Journal of Management*, 16: 461-509.
- Hoskisson, R. E., Hitt, M. A., Johnson, R. A., & Moesel, D. D. 1993. Construct validity of an objective (entropy) categorical measure of diversification. *Strategic Management Journal*, 14: 215-235.
- Jaffe, A. B. 1986. Technological opportunity and spillovers of R&D: Evidence from firms' patents, profits, and market value. *American Economic Review*, 76: 984-1001.
- Jacquemin, A., & Berry, C. 1979. Entropy measure of diversification and corporate growth. *Journal of Industrial Economics*, 27: 359-379.
- Karnani, A., & Wernerfelt, B. 1985. Multiple point competition. *Strategic Management Journal*, 6: 87-96.
- Kim, D., & Kogut, B. 1996. Technological platforms and diversification. *Organization Science*, 7: 283-301.
- Klein, P. G. 2005. The make-or-buy decision: Lessons from empirical studies. In C. Ménard & M. Shirley (Eds.), *Handbook of new institutional economics*. Dordrecht, the Netherlands: Springer.
- Kochar, R., & Hitt, M. A. 1998. Linking corporate strategy to capital structure: Diversification strategy, type and source of financing. *Strategic Management Journal*, 19: 601-610.
- Kwoka, J. E., & Ravenscraft, D. J. 1986. Cooperation v. rivalry: Price-cost margins by line of business. *Economica*, 53(211): 351-363.
- Lachmann, L. M. 1956. *Capital and its structure*. London: Bell.
- Larsson, R., & Finkelstein, S. 1999. Integrating strategic, organizational and human resource perspectives on mergers and acquisitions. *Organization Science*, 10: 1-26
- Larsen, K., & Meliciani, V. 2000. The importance of technology-based intersectoral linkages for market share dynamics. *Review of World Economics*, 136: 702-723.
- Lemelin, A. 1982. Relatedness in the patterns of interindustry diversification. *Review of Economics and Statistics*, 64: 646-657.
- Levin, C. R. Klevorick, A. K., Nelson, R. N., Winter, S. G., Gilbert, R., & Griliches, Z. 1987. Appropriating the returns from industrial research and development. *Brookings Papers on Economic Activity*, 3: 783-831
- Lien, L. B., & Klein, P. G. 2006. Relatedness and acquirer performance. In C. L. Cooper & S. Finkelstein (Eds.), *Advances in mergers and acquisitions*, Vol. 5. Amsterdam: Elsevier.
- March, J. G. 1991. Exploration and exploitation in organizational learning. *Organization Science*, 2(1): 71-87.
- Markides, C. C., & Williamson, P. J. 1994. Related diversification, core competence and corporate performance. *Strategic Management Journal*, 15: 149-165.
- Markides, C. C., & Williamson, P. J. 1996. Corporate diversification and organizational structure: A resource-based view. *Academy of Management Journal*, 39: 340-367.
- McFadden, D., & Manski, C. F. 1981. Alternative estimators and sample designs for discrete choice analysis. In D. McFadden & C. F. Manski (Eds.), *Structural analysis of discrete data with econometric applications*. Cambridge, MA: MIT Press.
- Milgrom, P., & Roberts, J. 1992. *Economics, organization and management*. Englewood Cliffs, NJ: Prentice-Hall.
- Miller, D. J. 2004. Firms' technological resources and the performance effects of diversification: A longitudinal study. *Strategic Management Journal*, 25: 1097-1119.
- Miller, D. J. 2006. Technological diversity, related diversification, and firm performance. *Strategic Management Journal*, 27: 601-619.
- Montgomery, C. A., & Hariharan, S. 1991. Diversified expansion by large established firms. *Journal of Economic Behavior and Organization*, 15(1): 71-89.
- Myers, S. C., & Majluf, N. S. 1984. Corporate financing and investment decisions when firms have information that investors do not have. *Journal of Financial Economics*, 13: 187-221.
- Nelson, R. R., & Winter, S. G. 1982. *An evolutionary theory of economic change*. Cambridge, MA: Belknap.
- Palich, L. E., Cardinal, L. B., & Miller, C. C. 2000. Curvilinearity in the diversification-performance linkage: An examination of over three decades of research. *Strategic Management Journal*, 21: 155-174.
- Penrose, E. 1959. *The theory of the growth of the firm*. Oxford, UK: Blackwell.
- Phillips, O. R., & Mason, C. F. 1996. Market regulation and multimarket rivalry. *Rand Journal of Economics*, 27: 596-617.
- Piscitello, L. 2000. Relatedness and coherence in technological and product diversification of the world's largest firms. *Structural Change and Economic Dynamics*, 11: 295-315.

- Porter, M. E. 1980. *Competitive strategy*. New York: Free Press.
- Porter, M. E. 1987. From competitive advantage to corporate strategy. *Harvard Business Review*, 65(3): 43-59.
- Prahalad, C. K., & Bettis, R. A. 1986. The dominant logic: A new linkage between diversity and performance. *Strategic Management Journal*, 7: 485-501.
- Rajan, R., Servaes, H., & Zingales, L. 2000. The cost of diversity: The diversification discount and inefficient investment. *Journal of Finance*, 55: 35-80.
- Reed, R., & Luffman, G. A. 1986. Diversification: The growing confusion. *Strategic Management Journal*, 7: 29-35.
- Richardson, G. B. 1972. The organisation of industry. *Economic Journal*, 82(327): 883-896.
- Robins, J., & Wiersema, M. F. 1995. A resource-based approach to the multibusiness firm: Empirical analysis of portfolio interrelationships and corporate financial performance. *Strategic Management Journal*, 16(4): 277-299.
- Robins, J. A., & Wiersema, M. F. 2003. The measurement of corporate portfolio strategy: Analysis of the content validity of related diversification indexes. *Strategic Management Journal*, 24: 39-59.
- Rumelt, R. P. 1974. *Strategy, structure, and economic performance*. Cambridge, MA: Harvard University Press.
- Rumelt, R. P. 1982. Diversification strategy and profitability. *Strategic Management Journal*, 3: 359-369.
- Salinger, M. A. 1984. Tobin's Q, unionization, and the concentration-profits relationship. *Rand Journal of Economics*, 15: 159-170.
- Santalo, J., & Becerra, M. 2008. Competition from specialized firms and the diversification-performance linkage. *Journal of Finance*, 68: 851-883.
- Scharfstein, D. S., & Stein, J. C. 1990. Herd behavior and investment. *American Economic Review*, 80: 465-479.
- Scharfstein, D. S., & Stein, J. C. 2000. The dark side of internal capital markets: Divisional rent-seeking and inefficient investment. *Journal of Finance*, 55: 2537-2564.
- Schmalensee, R. 1989. Inter-industry studies of structure and performance. In R. Schmalensee & R. D. Willig (Eds.), *Handbook of industrial organization*. New York: North Holland.
- SIC Manual. 1987. Executive Office of the President, Office of Management and Budget.
- Silverman, B. S. 1999. Technological resources and the direction of corporate diversification: Toward an integration of the resource-based view and transaction cost economics. *Management Science*, 45: 1109-1124.
- Silverman, B. S., Nickerson, J. A., & Freeman, J. B. 1997. Profitability, transactional alignment, and organizational mortality in the U.S. trucking industry. *Strategic Management Journal*, 18: 31-52.
- Shelanski, H. A., & Klein, P. G. 1995. Empirical research in transaction cost economics: A review and assessment. *Journal of Law Economics and Organization*, 11: 335-361.
- Shin, H., & Stulz, R. M. 1998. Are internal capital markets efficient? *Quarterly Journal of Economics*, 113: 531-552.
- Stein, J. C. 1997. Internal capital markets and the competition for corporate resources. *Journal of Finance*, 52: 111-133.
- Teece, D. J. 1980. Economies of scope and the scope of the enterprise. *Journal of Economic Behavior and Organization* 1: 223-247.
- Teece, D. J. 1982. Towards an economic theory of the multiproduct firm. *Journal of Economic Behavior and Organization*, 3: 39-63.
- Teece, D.J. 1986. Profiting from technological innovation. *Research Policy*, 15: 285-306.
- Teece, D. J., Rumelt, R., Dosi, G., & Winter, S. 1994. Understanding corporate coherence: Theory and evidence. *Journal of Economic Behavior and Organization*, 23: 1-30.
- Villalonga, B. 2004. Does diversification cause the diversification discount? *Financial Management*, 33(2): 5-27.
- Voigt, C. W. 1993. *The Trinet large establishment data*. Unpublished manuscript, Anderson Graduate School of Management, University of California, Los Angeles.
- Williamson, O. E. 1975. *Markets and hierarchies: Analysis and antitrust implications*. New York: Free Press.
- Williamson, O. E. 1985. *The economic institutions of capitalism: Firms, markets, relational contracting*. New York: Free Press.
- Winter, S. G. 1964. Economic natural selection and the theory of the firm. *Yale Economic Essays*, 4(1): 225-273.
- Winter, S. G. 1971. Satisficing, selection, and the innovating remnant. *Quarterly Journal of Economics*, 85: 237-261.
- Yip, G. S. 1982. Diversification entry: Internal development versus acquisition. *Strategic Management Journal*, 3: 331-345.
- Zenger, T. R., & Marshall, C. R. 2000. Determinants of incentive intensity in group-based rewards. *Academy of Management Journal*, 43: 149-163.