CAN THE SURVIVOR PRINCIPLE SURVIVE DIVERSIFICATION?

Lasse B. Lien

Department of Strategy and Management
Norwegian School of Economics and Business Administration
Breiviksveien 40, N-5045 Bergen, Norway
lasse.lien@nhh.no

Peter G. Klein

Division of Applied Social Sciences
University of Missouri
135 Mumford Hall, Columbia, MO 65211 USA
and
Department of Strategy and Management
Norwegian School of Economics and Business Administration
Breiviksveien 40, N-5045 Bergen, Norway
pklein@missouri.edu

This version: 16 November 2010

Abstract: The survivor principle holds that the competitive process weeds out inefficient firms, so that hypotheses about efficient behavior can be tested by observing how firms actually behave. This principle underlies a large body of empirical work in strategy, economics, and management. But do competitive markets actually display what is efficient? Is the survivor principle reliable? We evaluate the survivor principle in the context of corporate diversification, showing that survivor-based measures of inter-industry relatedness are good predictors of firm’s decisions to exit particular lines of business, even when controlling for other firm and industry characteristics that affect the decision to withdraw from one industry or another. We argue further that his relatedness measure captures an important aspect of economic efficiency, not simply firms’ desire for legitimacy by imitation, or attempts to temper multi-market competition. Hence confidence in the survivor principle is warranted, at least in this context.

Key words: diversification, exit, relatedness, survivor principle

We thank Nicolai Foss, Harvey James, Olav Kvitastein, Jackson Nickerson, Brian Silverman, Anne Yvrande-Billon, the editor, two anonymous reviewers, and participants at the Nordic Workshop on Transaction Cost Economics, Strategic Management Society meeting, and International Society for New Institutional Economics conference for helpful comments. Don Hatfield and Carl Voigt generously shared the AGSM/Trinet database. The authors are responsible for any remaining errors and obscurities.
Empirical work in organizational economics, corporate finance, and strategic management often begins with the assumption that the competitive selection process works well, so that surviving firms tend to display efficient behavior or characteristics. To test whether a particular behavior is efficient, we simply look to see whether this behavior dominates within a population of firms in competitive markets. For example, to test whether specific investments lead to vertical integration, we pick a sample of firms and examine the relationship between vertical integration and a measure of asset specificity (Klein, 2005). To find out if the incentive effects of performance-based pay outweigh the losses from inefficient risk sharing, we study the correlation between the use of performance-based pay and firm or industry characteristics (Prendergast, 1999). In other words, to see what strategies or structures work well under particular circumstances we look at actual behavior, assuming that the competitive process will ensure that the efficient choices dominate.

The assumption that inefficient forms of behavior are selected out is called the *survivor principle*. While the name was coined by Stigler (1968), the idea is usually attributed to Alchian (1950) and Friedman (1953). Alchian argued that even though theories about rational decision makers making “optimal” choices are clearly unrealistic, the predictions of such theories need not be. The quest for profit, combined with competitive selection forces, ensures that the average firm will tend to behave like those described by theories of rational behavior (Alchian, 1950). Friedman (1953: 22), defending the profit-maximization hypothesis, puts it this way:

> Unless the behavior of businessmen in some way or other approximated behavior consistent with the maximization of returns, it seems unlikely that they would remain in business for long. Let the apparent immediate determinant of business behavior be anything at all—habitual reaction, random choice, or whatnot. Whenever this determinant happens to lead to behavior consistent with rational and informed maximization of re-

---

1 Some trace the concept back to Harrod (1938). Another early contributor is Enke (1951).
turns, the business will prosper and acquire resources with which to expand; whenever it
does not, the business will tend to lose resources and can be kept in existence only by the
addition of resources from outside. The process of “natural selection” thus helps to valid-
date the [maximization] hypothesis or, rather, given natural selection, acceptance of the
hypothesis can be based largely on the judgment that it summarizes appropriately the
conditions for survival.

Note that the general claim is that surviving firms will behave efficiently, not necessarily that this be-
havior is particularly well described by neoclassical economics. Thus, while transaction cost economists
may claim that the efficiency calculus of neoclassical economics gives insufficient consideration to
bounded rationality and opportunism (Williamson, 1975, 1985), and resource-based theorists may claim
that it downplays factor market imperfections (Barney, 1986; Dierickx and Cool, 1989; Wernerfelt,
1984); they may still accept a general version of the survivor principle. Williamson (1988: 174), for ex-
ample, notes that empirical research in transaction cost economics “relies in a general, background way
on the efficacy of competition to perform a sort between more and less efficient modes and to shift re-
sources in favor of the former.”

Some version of the survivor principle underlies a substantial body of empirical research in strategy,
economics and management. Yet, as elaborated below, several authors remain skeptical as to whether the
competitive process can perform the screening function ascribed to it by the survivor principle. Given its
widespread and controversial use, it is surprising that the survivor principle itself has not been subject to
empirical testing. Our intention is to conduct such a test in the context of exit decisions by diversified
firms.

We begin by assuming that which industries a diversified firm combines has consequences for the ef-
ficiency of the firm. If the survivor principle holds in the context of corporate diversification, we can fur-
ther assume that those pairs of industries that are most frequently combined in firms will on average
represent more efficient combinations than those pairs of industries that are rarely combined. In other
words, the behavior of competitive firms can show us which combinations are efficient. If so, a diversified firm should be significantly more likely to exit those industries that this “survivor logic” identifies as a poor match with other businesses in the portfolio, compared to those identified as representing a good match.

We test this prediction using detailed data on firm’s business portfolios from the AGSM/Trinet database for the early 1980s to construct a survivor-based measure of relatedness and ask how well this measure predicts the exit decisions of diversified firms. As we show below, the survivor-based measure of relatedness is a strong predictor of exit, even when controlling for other firm and industry characteristics that might affect the decision to withdraw from a particular industry. We provide further evidence that the survivor-based relatedness measure captures an important aspect of economic efficiency, not simply firms’ desire for legitimacy by imitation, or attempts to temper multi-market competition. During that period, then, the competitive process did tend to filter out inappropriate business combinations.

THE SURVIVOR PRINCIPLE

When Alchian articulated the survivor principle in 1950, mainstream economic theory was widely criticized for its unrealistic representation of human decision making. Critics argued that economic actors had neither the information nor the processing capabilities assumed in standard models (Simon, 1947) and did not think in terms of marginal analysis (Lester, 1946). Alchian countered that even if neoclassical models do not describe how people actually behave, they can still make accurate predictions, for two reasons. First, firms making negative profits must either take corrective measures or lose resources and ultimately go out of business, while firms earning profits will acquire resources and grow. Second, the profit motive provides strong incentives for less successful firms to imitate the more successful firms.
Together, Alchian argued, these forces ensured that surviving firms in competitive markets would appear “as if” they were behaving as described by neoclassical economic theory.²

Friedman (1953) extended Alchian’s ideas in two directions. First, while Alchian tolerated unrealistic behavioral assumptions as needed for describing a competitive selection process, Friedman treated unrealistic assumptions as a virtue. A good theory, according to Friedman, is a theory that explains much by little, generating good predictive accuracy based on simplifying assumptions that remove the clutter and detail of the real world. He praised neoclassical price theory as useful, not because its assumptions are realistic, but because the forces described by Alchian would ensure predictive accuracy. Second, Friedman made bolder claims than Alchian about the accuracy of these predictions. While Alchian thought the competitive selection process merely chose the best of the tested alternatives (survival of the fitter), Friedman claimed that competition would yield “optimal” behavior (survival of the fittest).

Alchian’s and Friedman’s arguments underlie both theory and empirical work. Theoretically, they are taken to justify efficiency explanations for institutions and behavior, even when we know little about the underlying mechanisms (Dow, 1987). Such functionalist explanations dominate organizational economics and strategy. The other use relates to theory testing. If the survivor principle holds, then the behavior of a randomly selected group of competitive firms can be taken as efficient. Empirical studies need not deal with efficiency directly, but can focus instead on the reduced-form relationship between exogenous characteristics and behavior.

This empirical strategy is widely adopted, though rarely stated explicitly. In transaction cost economics, the hypothesis that vertical integration is more efficient than market governance in the presence of

---

² Use of the survivor principle in social-science research is often seen as an application of evolutionary biology, though Herbert Spencer (1864) developed his account of social evolution two years before Darwin’s *Origin of Species*. Spencer (1884) later applied the evolutionary principle to competition among societies, echoing the reasoning in Burke’s (1790) critique of the French Revolution. John Stuart Mill’s *On Liberty* (1859) includes a survivor-based discussion of freedom of speech; only by having access to alternative opinions can people pit alternative arguments against each other as they work toward the truth.
asset specificity is usually tested by seeing whether firms actually integrate when asset specificity is high (see Shelanski and Klein, 1995; David and Han, 2004; and Macher and Richman, 2008 for surveys of a large empirical literature). In agency theory, hypotheses about the relative efficiency of alternative incentive contracts are often examined by identifying which contracts firms actually use (Anderson, 1985; Eisenhardt, 1985; Zenger and Marshall 2000). In the diversification literature in strategic management, theories about efficient diversification are typically examined by observing real-world diversification patterns (Montgomery and Hariharan, 1991; Farjon, 1994; Silverman, 1999). The survivor principle is also at work in empirical studies in industrial organization, property rights theory, corporate finance, and marketing. Clearly it is central to empirical work in the study of organizations and their behavior.

From the outset, however, critics have questioned the survivor principle’s role in theory testing (Penrose, 1952). Is the competitive selection process fast and precise enough to justify the assumption that surviving firms are efficient? Friedman’s optimizing version also suffers from the problem of sufficient variation—only behaviors that are part of the initial set of behaviors can be chosen as efficient survivors (Nelson and Winter, 1982). Entrepreneurial experimentation and learning may bring forth the optimal behavior, but unless incremental change is continually beneficial—i.e., there is no local optimum next to the global optimum—the optimal solution may not be reached through small steps (Elster, 1989). And if

---

3 Some newer papers in the empirical transaction cost literature do not assume the SP, but rather employ a two-stage procedure in which in which the relationship between transactional characteristics and governance structure is endogenously chosen in the first stage, then used to explain performance in the second stage. Silverman, Nickerson and Freeman (1997) and Nickerson and Silverman (2003), for example, show that transaction cost efficiency is positively correlated with firm survival in the for hire trucking industry, while Argyres and Bigelow (2007) and Bigelow and Argyres (2008) examine outsourcing arrangements in the U.S. automobile industry and find that transactions that are appropriately aligned tend to last longer than inappropriately organized ones.

4 A related set of objections deals with the more basic goals of social science. If explanation, not prediction, is task of economic theory, then the survivor principle—and the functionalist explanations it endorses—directs our attention away from theories that provide true accounts of the relevant causal processes. For more, see the controversy over Friedman’s (1953) defense of unrealistic assumptions (Boland, 1979; Blaug, 1980; Caldwell, 1980; Musgrave, 1981; Mäki, 1994).
experimentation and learning are slow and gradual, then inefficient behaviors can persist for long periods, rendering the optimizing version of the survivor principle untenable.

Alchian’s comparative-efficiency version of the survivor principle—the version relied upon more commonly in empirical work—is not vulnerable to this critique. His argument is that selection will operate on the tested solutions only, and bring about a situation where the best of these dominates. This does not imply optimality: “Positive profits accrue to those who are better than their competitors, even if the participants are ignorant, intelligent, skillful, etc. . . . As in a race, the award goes to the relatively fastest, even if all competitors loaf” (Alchian, 1950: 213). However, this more modest version is not immune to all criticisms. 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; the populations we observe today may be dominated by the solutions that were efficient yesterday.

A related objection is that the competitive process selects for overall firm performance, not the individual decisions that determine performance (Elster, 1989). Because performance results from a cluster of choices, a firm may survive even if some of its managers’ decisions are inefficient, as long as other decisions are sufficiently efficient to ensure survival. Moreover, there is a feedback mechanism between market selection and environmental conditions (Hodgson, 1993). Early entrants may experience first-mover advantages, which dissipate as rival firms follow suit, such that the early mover’s formerly profitable behavior becomes unprofitable. If such scenarios occur often, then populations will appear to be dominated by comparatively inefficient behaviors, even if these behaviors were efficient when initially undertaken.

---

5 Alchian’s version is the one implicitly used in most empirical work because hypotheses are typically formulated as comparative statements about efficiency, for example between hierarchical and market governance, related and unrelated diversification, fixed or variable compensation, and the like. Empirical tests of such comparative statements do not rely on the optimizing version of the survivor principle, but only that relatively more efficient outcomes will tend to be observed.
Finally, as emphasized by Masten, Meehan, and Snyder (1991) a reduced-form model may be consistent with several, possibly contradictory, structural models. For example, asset specificity and hierarchical governance may go together not because asset specificity increases the cost of market governance by exposing parties to potential holdups, as the theory maintains, but because asset specificity somehow reduces the cost of hierarchy (Ghoshal and Moran, 1996). Reliance on the survivor principle does not alleviate this concern.6

Given these concerns, is reliance on the survival principle reasonable? After all, trying to measure efficiency directly poses problems as well (Benston, 1985; Fisher and McGowan, 1983). To address this issue, we subject the survivor principle itself to empirical scrutiny. A logical first step in doing so is to examine the assumption that decisions or behaviors that occur frequently in a population of competitive firms are on average more efficient than those that occur rarely. Here we offer such a test in the context of corporate diversification.

A SURVIVOR-BASED MEASURE OF RELATEDNESS

As noted above, we assume that which industries a diversified firm includes in its portfolio will affect firm performance. We refer to the performance effects of these combinations as the relatedness of a given industry to other industries in the firm’s portfolio.7 We make no assumptions about exactly what causes

6 These criticisms are rarely addressed in the empirical literature. An exception is Williamson (1988: 174), who acknowledges that the process of transaction cost economizing is not automatic. The claim that transactions and governance structures are efficiently aligned, he says, “seems plausible, especially if the relevant outcomes are those that appear over intervals of five and ten years rather than in the very near term. This intuition would nevertheless benefit from a more fully developed theory of the selection process.” Despite an extensive literature in evolutionary game theory (Banerjee and Weibull, 1996; Fudenberg and Levine, 1999), there is little consensus about how selection mechanisms work, and to what extent they can be relied upon in empirical studies of efficient behavior.

7 Relatedness has primarily been studied at the inter-industry level, but since firms in any industry differ in their resources there may be a relatedness component that is firm specific (see, for example, Silverman, 1999). In line with most of the existing work, we focus on the inter-industry component of relatedness (i.e. not the firm specific component). Note that research on corporate diversification has repeatedly documented the existence of stable and
these performance effects, but instead we assume—following the survivor principle—that the decisions made by competitive firms can reveal the relatedness between any given pair of industries. Hence we use the term differently from Rumelt (1974), who defined relatedness ex ante based on characteristics of a firm’s portfolio.\(^8\)

The fundamental premise of this survivor-based approach to relatedness is thus that related industries are more frequently combined in firms than unrelated industries. More 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 call this difference a survivor-based measure of the relatedness between a pair of industries. This method was originally proposed and developed by Teece, Rumelt, Dosi, and Winter (1994). For any pair of industries \(i\) and \(j\) the procedure generates a standardized measure we call “survivor relatedness” or \(SR_{ij}\) reflecting the difference between the likelihood of observing firms active in both \(i\) and \(j\) and what we would observe if firms chose their industries at random. The measure adjusts for industry size such that it can be compared consistently across industry pairs. Hence \(SR_{ij}\) can be used to establish a survivor-based measure of the relatedness of a given business to the other businesses in a corporate portfolio.

---

\(^8\) Rumelt (1974) focuses on three measures of relatedness, the proportion of a firm’s revenue attributable to its largest single business (specialization ratio), the proportion of a firm’s revenue attributable to its largest group of related businesses (related ratio), and the proportion of a firm’s revenue arising from all byproducts, intermediate products, and end products of a vertically integrated sequence of processing activities (vertical ratio).
Computing relatedness

To compute $SR_{ij}$, let $K$ be the universe of diversified firms, and $I$ be the universe of these firms’ industries. Let $C_{ik} = 1$ if firm $k$ is active in industry $i$. The number of industries in firm $k$’s portfolio is $m_k = \Sigma C_{ik}$ and the number of diversified firms in industry $i$ is $n_i = \Sigma_k C_{ik}$. Next, let $J_{ij}$ be the number of diversified firms active in both industries $i$ and $j$, such that $J_{ij} = \Sigma_k C_{ik}C_{jk}$. 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 $j$, we compare $J_{ij}$ to the number of expected combinations if diversification patterns were random.

To calculate the number of expected combinations we construct a sample of size $n_i$ drawn without replacement from a population of $K$ firms. Those firms are considered active in industry $i$. A second independent sample of size $n_j$ is then drawn from the population of $K$ firms. Those firms 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:

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

(1)

The mean and variance of $X_{ij}$ are:

$$\mu_{ij} = E(X_{ij}) = \frac{n_in_j}{K},$$

(2)

$$\sigma^2 = \mu_{ij}\left(1-\frac{n_j}{K}\right)\left(\frac{K}{K-1}\right).$$

(3)

The expected number of combinations under random diversification is $\mu_{ij}$, and we can now construct a standardized measure of the difference between actual and expected combinations in the following way:

---

9 See [reference withheld] for further details.
Note that the measure adjusts for the number of firms in each industry, as larger industries have more potential matches to choose from.10

We then use the measure $SR_{ij}$ to establish a survivor-based measure of the relatedness of a given business (or potential business) to the other businesses in a corporate portfolio. We call this measure portfolio match or $PM_i$. Again the procedure is based on Teece et al. (1994). Assume a diversified firm participates in $m$ industries. Its business in industry $i$ has sales of $s_i$ and survivor-based relatedness $SR_{ij}$ with industry $j$.

We first rank the survivor-based relatedness $SR_{ij}$ for all pairs between industry $i$ and other industries in the firm’s portfolio. We call the two industries with the highest score the “neighboring” businesses. Let $\lambda_{ij} = 1$ for each business $j$ defined in this way as a neighbor to business $i$ and $\lambda_{ij} = 0$ for the other $j$ businesses. The weighted average survivor-based relatedness of the portfolio neighbors to business $i$ is then defined by

$$PM_i = \frac{\sum_{j \neq i} SR_{ij} s_j \lambda_{ij}}{\sum_{j \neq i} s_j \lambda_{ij}}.$$ (5)

This ratio can be considered a survivor-based, parent-to-subsidiary measure of relatedness because it is constructed under the assumption that we can infer the efficiency of various combinations of industries by observing the diversification decisions made by competitive firms. In other words, $PM_i$ is essentially the application of $SR_{ij}$ to characterize the relationship between the subsidiary $i$ and the most relevant other subsidiaries in the parent’s portfolio.

**Hypotheses about relatedness and performance**

---

10 We also control for industry structure, using a concentration measure in our regressions below.
Assuming stochastic variation in firms’ entry decisions and in environmental conditions, some diversified firms will contain combinations the survivor-based logic designates as inefficient. The survivor principle implies that firms are more likely to exit industries with low survivor-based relatedness scores than those with high scores. Continuing to operate in low-scoring business lines reduces firm performance, for several reasons, such as the loss of scope economies and additional complexity and governance costs to the firm (Prahalad and Bettis, 1986). We are agnostic here about the precise mechanisms. Moreover, a firm’s unrelated businesses will tend to be attractive targets to firms with complementary businesses in their portfolios (Goold, Campell, and Alexander, 1994). For these reasons we expect the probability that a firm will exit a given line of business—by liquidation or divestiture—to decrease with the relatedness of that business unit to the firm’s other businesses. More specifically, we post the following hypotheses:

H1: The lower a business scores on PM, the more likely the parent will exit that business, other things equal.

Rejecting H1 implies either that (a) firm performance is independent of the industry combinations within a firm’s portfolio, or (b) firms’ diversification decisions provide no information about the efficiency of various business combinations. The former is unlikely, so our analysis constitutes an empirical test of the latter—i.e., a test of the survivor principle itself.\footnote{Admittedly, evidence on the performance effects of portfolio composition is mixed (Hoskisson and Hitt, 1990; Ramanujam and Varadarajan, 1989; Robins and Wiersma, 1995), but this is likely because relatedness is difficult to measure, not because portfolio composition is unrelated to performance. A particularly difficult problem is that the causes of relatedness vary across situations, making relatedness using categorical, functional measures difficult to handle in large-sample research. Our approach avoids this problem because we do not impose a particular view of relatedness on the data. Instead we let the data tell us what is related to what.}

**Alternative interpretations**
Rejecting H1 would constitute strong evidence against the survivor principle. But firms’ diversification decisions might tend to be clustered for reasons other than efficiency. Institutional theory, for example, suggests that firms seek legitimacy in the eyes of various important constituents and stakeholders to secure support and critical resources (DiMaggio and Powell, 1983, 1991). A key mechanism used to obtain legitimacy is mimicry, leading to institutional isomorphism—organizations facing similar environmental conditions will tend to adopt the same structures (Meyer and Rowan, 1977; DiMaggio and Powell, 1983). In other words, the quest for legitimacy leads firms to adopt structures that are considered appropriate and rational, resulting in structural convergence. Of course, if these institutional pressures always promote structures that are efficient, the SP and the institutional alternative become observationally equivalent, at least in large-sample research. However, institutional theory suggests that convergence can occur even when the shared structure is not economically efficient (Scott 1995; Zucker, 1987).

There is a substantial literature on corporate diversification from an institutional or neo-institutional perspective (e.g. Davis, Diekman and Tinsley, 1994; Fligstein, 1991; Lounsbury, 2004; Zuckerman, 2000; Phillips and Zuckerman, 2001). Of particular interest is the refocusing movement of the early and mid 1980s, exactly the period we study here (Bhide, 1990; Shleifer and Vishny, 1991). The institutional explanation is that firms require legitimacy from financial-market participants, particularly financial analysts, institutional investors, and other shareholders. As these big players began to frown upon unrelated diversification in the early 1980s, they successfully pressured corporate decision makers to refocus (Fligstein, 1991; Zuckerman, 2000; Phillips and Zuckerman, 2001). The pressure on corporate managers operated through analysts’ willingness to cover and recommend stocks, and through investors’ pricing and willingness to participate in emissions.

To test whether institutional pressures that are weakly linked to efficiency may account for the pattern hypothesized in H1, we note that our sample period is one in which these institutional pressures were

---

12 We thank two anonymous reviewers for clarifying our thinking on this point.
presumed to be particularly strong, and that they should hit publicly listed firms first and hardest. If institutional isomorphism, not efficiency, is driving behavior, then survivor-based relatedness should be a good predictor of exit for publicly listed firms, but less so for unlisted firms. Assuming that institutional pressure and not efficiency explains the pattern in H1, we post the following hypothesis:

H2: Other things equal, listed firms are more likely than unlisted firms to exit businesses scoring low on PM.

Another possible explanation for the pattern described in H1 is mutual forbearance. The mutual forbearance hypothesis, first proposed by Edwards (1955), suggests that multipoint competitors will refrain from aggressively attacking each other. Frequent contact across markets allows firms to respond to aggressive actions by multipoint rivals in other markets in which the firms compete, and the threat of such retaliation limits competition in the focal market (Karnani and Wernerfelt, 1985). Previous studies associate multipoint competition with higher prices (Feinberg, 1985; Evans and Kessides, 1994; Gimeno and Woo, 1996; Gimeno, 1999), higher profits (Barnett et al., 1994; Phillips and Mason, 1996), and lower exit rates (Barnett, 1993; Baum and Korn, 1996; Baum and Korn, 1999; Baum, 1999).

If firms prefer multipoint competition, to limit the aggressive behavior of rivals, then they might refrain from exiting a weak position in one industry to protect gains from mutual forbearance in another. In other words, firms retain underperforming businesses, not because those businesses are part of an efficient portfolio, but to retain some other kind of strategic advantage. Note that the survivor-based measure of industry relatedness on which PM is constructed counts the frequencies of multimarket contact across industries, and might therefore be reflecting motives related to mutual forbearance, not efficiency. If so, we cannot interpret support for H1 as an argument about efficient industry combinations.

To examine this problem we note that mutual forbearance requires some minimum level of concentration in the relevant markets to be a plausible motive for portfolio choices. In fragmented markets, exist-

---

13 In our regressions we control for firm size and initial diversity, to account for the fact that large and highly visible private firms could also be subject to institutional pressures.
ing or potential competition from single-business firms are likely to force multimarket firms into competing vigorously. If our portfolio match measure $PM$ reflects mutual forbearance, we would expect the following:

**H3:** $PM$ is a better predictor of exit in the presence of industry concentration.

**METHODS**

Our approach proceeds in two steps. The first is to develop the survivor-based measure of industry relatedness $SR_{ij}$ for each pair of industries that is combined in diversified firms, to then construct our firm-level independent variable $PM$. The second step is to test $H1–H3$. As the data, samples, variables, and methods differ between the two steps, we discuss the methodological issues separately for each.

Our survivor-based relatedness measure is based on the AGSM/Trinet Large Establishment Database (Trinet). Trinet contains biannual records of all U.S. establishments with more than 20 employees from 1979 to 1989, including data on 4-digit SIC code, corporate ownership, and sales.\(^{14}\) By aggregating the establishments for each parent in each 4-digit SIC code, the 4-digit SIC codes for each parent, and the parents for each 4-digit SIC industry, we get a comprehensive picture of diversification patterns in the U.S. economy.\(^{15}\) While several decades old, the Trinet database remains an important source of U.S. business information.

---

\(^{14}\) Trinet also includes foreign establishments in the U.S. Because of changes in the parent coding in the Trinet database in 1979, and changes in the SIC classification scheme in 1987, only data from the years 1981, 1983 and 1985 are directly comparable.

The Trinet database contains some 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 0.893 which indicates that the sales data in Trinet are of acceptable quality. The problem of imputed sales data is likely to be largest for the smallest firms in the Trinet database (which are omitted from our sample).

\(^{15}\) Comparison with the Census of Manufacturers indicates that Trinet contains 95 percent of all establishments it should (Voigt, 1993) and that omissions are most likely for small firms (which are less likely to be diversified).
ness-level activity. Compustat’s Business Information File, available annually since the early 1980s, contains more recent data but provides financial operation only for the firm’s largest business segments (defined as those contributing 10 percent or more to total sales), and nothing on operating units within segments. As noted above, institutional pressures for refocusing were particularly strong during the 1980s, making the Trinet database particularly useful for our purposes. The U.S. Census Bureau has recently created a new establishment-level database, the Business Information Tracking Series (BITS), that is comparable to Trinet (see Villalonga, 2004).

The industry match \( SR_{ij} \) is calculated from the Trinet files of 1981 using all firms active in two or more 4-digit SIC-codes. After deleting single-business firms, government-owned firms, and firms in non-profit sectors, we have a sample of 13,164 diversified firms, active in 929 different industries and covering 57,647 individual business units. Of the 431,056 possible industry pairs, 122,105 were observed. The measure of \( SR_{ij} \) between the observed industry pairs ranged from \(-7.97\) to 93.55 with a mean of 4.33 and a standard deviation of 5.06.\(^{16}\)

Based on the estimates of survivor relatedness, we calculate the portfolio matches \( PM \) following the procedures described above.

Sample

H1 states that if the survivor principle holds, businesses with low \( PM \) scores (i.e., businesses unrelated to the other businesses in a corporate portfolio) are more likely to be exited than businesses with high \( PM \) scores. H2 states that the pattern described by H1 will be substantially stronger for listed firms.

\(^{16}\) The most highly related of all four-digit SIC industries are 8011, Offices of Physicians, and 8081, Outpatient Care Facilities (\( SR_{ij} = 93.55 \)). Another highly related pair of industries is 4411, Deep Sea Foreign Transportation, and 4723, Freight Transportation Arrangement (\( SR_{ij} = 82.96 \)). Note that these are in different two-digit SIC groups, and would therefore be classified as unrelated according to measures based on distances between SIC codes. Additional examples are 6311 and 6321, Life Insurance and Accident and Health Insurance (\( SR_{ij} = 77.71 \)); 2517 and 3672, Wood Cabinets for Television, Radio, Phonography and Cathode Ray Television Tubes (\( SR_{ij} =77.54 \)); and 4832 and 4833, Radio Broadcasting and Television Broadcasting (\( SR_{ij} = 70.60 \)).
than for unlisted firms. H3 states that the pattern described by H1 will be substantially stronger for exits from concentrated industries than from fragmented industries. To examine these hypotheses, we begin with all 6,416 firms in the Trinet database that were present in 3 or more 4-digit SIC codes in 1989. (The requirement that a firm should be in 3 or more industries is necessary to ensure that $PM$ is calculated consistently across firms.) We then remove firms that were sold or liquidated in their entirety between 1981 and 1985. This is because such actions do not reveal information about the merits of combining different industries inside a firm, while exiting some businesses and keeping others do.\textsuperscript{17} This reduces the number of parent firms to 5,660. These 5,560 firms operated 39,434 businesses in 924 4-digit SIC-codes in 1981. From these 31 cases were lost due to missing data, bringing the sample down to 39,403. Of these 39,403 businesses 8,148 were exited by 1985 while they remained in 31,219. This approximates a population level dataset of selective exit decisions made by diversified firms during the 1981–85 period.

It is crucial to note that the relatedness measure is based on data from the 1981 Trinet files, while exits are identified using the 1983 and 1985 files (i.e., exits after 1981). Because exits—along with none-entry—determines the key independent (relatedness) variable, and our dependent variable is the probability of exit, endogeneity may seem like a potential concern.\textsuperscript{18} However, past exit decisions cannot affect future exit decisions, as a firm cannot exit the same industry twice in the period covered in this study. Similarly, past decisions not to enter an industry cannot affect future exit decisions because a firm cannot exit an industry it has not entered. What we are investigating is essentially how much information nonentry

\textsuperscript{17} This introduces a survivor bias in our data, but one that is theoretically justified, because we want our data to shed light on the viability of particular industry combinations. Exit of an entire firm could have many causes other than the efficiency of its particular industry combinations. Selective exit (or not) by a continuing firm can be more readily interpreted as evidence on the merits of specific combinations. Fortunately, our conclusions remain the same if we leave these firms in the sample; the coefficients are generally smaller in magnitude, but the significance levels, and the overall interpretation, remain the same.

\textsuperscript{18} Strictly speaking the dependent variable is the log odds of exit, but we also report our results converted to effects on the probability of exit.
and exit decision by other firms at time $t_0$—as manifested in frequencies of industry combinations at time $t_0$—contains about the probability that a given firm exits a given industry by time $t_1$. If the survivor principle is a good description of empirical reality, the amount of information should be substantial. If it is not, the information will be poor.

A related concern is the amount of change in relatedness between $t_0$ and $t_1$. If the rate of change in relatedness due to technological innovation and changes in the competitive and regulatory environments is large, the noted information will be poor for reasons that do not necessarily invalidate the survivor principle. However, we find that the average relatedness between industries as measured by $SR_{ij}$ changes very little over the period covered in this study. The correlation between $SR_{ij}$ in 1981 and 1983 is 0.941, and between 1981 and 1985 is 0.895. If data were available for a longer period, we expect we would see more time-series variation.

To test H1–H3 we run a series of logistic regressions of the probability of exit on portfolio match and controls for industry- and parent-specific characteristics suggested by previous research to affect the exit decision. The general model is the following:

$$\text{Logit } Y = \beta_1 + \beta_2 \text{(industry growth)} + \beta_3 \text{(industry concentration)} + \beta_4 \text{(industry profitability)} + \beta_5 \text{(market share)} + \beta_6 \text{(parent size)} + \beta_7 \text{(parent diversity)} + \beta_8 \text{(PM)} + \epsilon. \quad (6)$$

The dependent variable is coded as follows. If a parent active in a 4-digit SIC code in 1981 has exited this business by 1985, the dependent variable gets a value of 1. If the parent is still active in the industry, the value assigned is 0. Both divestures and closures are thus considered to represent exit. \text{Logit } Y \text{ is the natural logarithm of the odds that an industry was exited:}

$$\ln \left[ \frac{p(Y = 1)}{1 - p(Y = 1)} \right] \quad (7)$$
The test of H2 involves running the same regression on separate samples of listed and un-listed firms. The test of H3 involves examining the interaction effect between industry concentration and survivor-based relatedness by adding a product term for industry concentration and PM.

**Industry-level independent variables**

To test H1 and H2 it is important to control for industry-level factors besides relatedness that affect exit. All things equal firms are less likely to exit an industry with high average profitability than an industry with low profitability. We add three industry-level variables presumed to affect the attractiveness of an industry: industry growth, industry concentration, and industry profitability. (Industry profitability is included to control for unspecified industry effects not captured by the two other industry-level variables.)

**Industry growth.** Industry growth is widely assumed to make an industry more attractive, as it allows firms to grow without having to attract customers away from rivals. Industry growth thus tends to soften competitive rivalry and raise average profitability. Such a relationship has been confirmed in several empirical studies (Kwoka and Ravenscraft, 1986; Salinger, 1984; Schmalensee, 1989). We therefore expect a negative relationship between industry growth and the probability of exit. We construct this variable as the percentage growth in industry sales 1981 and 1985, as reported in Trinet.

**Industry concentration.** Traditional industrial organization theory posits a positive relationship between industry concentration and industry profitability (Bain, 1956; Porter, 1980). Scale economies and other sources of market power, it is argued, reduce the threat from potential entrants, allowing incumbents more latitude to raise prices without inviting entry. Such a relationship was found in influential studies by Bain (1951), Weiss (1974), and Montgomery (1985), among others, but the relationship between concentration and profitability remains controversial (see Schmalensee, 1989, for a review). We expect a negative relationship between industry concentration and the probability that the industry will be exited. For concentration we use the average of the 4-firm concentration ratio for each industry in 1981 and 1983, based on Trinet.
Industry profitability. Industry profitability is affected by factors besides growth and concentration. To control for these factors we include in our regressions the average return on assets for each industry over the 1981–83 period, calculated from Compustat data. We calculate Industry profitability as total industry return divided by total industry assets. This provides a measure of the return on the average dollar invested in each industry. For this we use both Compustat’s industrial file and its business-segment file. We expect a negative relationship between industry profitability and the probability of exit.

Firm-level independent variables

Testing H1–H3 also requires that we control for firm-level factors other than the relatedness of the business in question. We control for four firm level characteristics thought to affect exit: market share, parent size, parent leverage, and parent liquidity.

Market share. Several studies document a positive relationship between market share and profitability (Gale, 1971; Robins and Wiersma, 1995; Sheperd, 1972), though the causal mechanism is unclear. Some explanations emphasize market power, while others focus on cost advantages from learning curve effects and economies of scale. Regardless of the underlying causal relation, we expect firms to be less likely to exit a business in which they hold a large market share. We construct market share as firm sales in industry $i$ as a percentage of total industry sales in 1981, based on the Trinet files.

Parent size. Parent size may be an indicator of market power and economies of scale. It is also correlated with the parent’s financial resources, tangible capital, and intangible assets. We thus expect a negative relationship between parent size and the probability of exit. We measure parent size as the logarithm of total sales in 1981, based on Trinet data.

Parent diversity. This variable records the number of 4-digit industries in the parent portfolio. It is included to control for the possibility that firms with large portfolios make exit decisions differently from firms with smaller portfolios, and for the general trend towards more focused portfolios occurring in our sample period. We expect Parent diversity to be positively related to the probability of exit.
**Portfolio match.** As described above, PM measures the sales-weighted average survivor-based relatedness of business \(i\) to the two closest neighboring businesses in parent \(k\)’s portfolio. We expect PM to be negatively related to the probability of exit.

Table 1 summarizes each variable definition, data sources, and predicted sign. Table 2 provides the means, standard deviations, and correlation coefficients for all independent variables.

[Tables 1 and 2 about here]

**RESULTS AND DISCUSSION**

Table 3 presents the results of two logistic regressions. Model 1 contains only the control variables. Model 2 adds PM, the sales-weighted relatedness of industry \(i\) to the two closest neighboring industries in the parent firm’s portfolio.

[Table 3 about here]

H1 predicts that PM is negatively related to the probability of exit, controlling for industry- and firm-specific characteristics that also affect exit. As seen from the second column of Table 3, the hypothesis is strongly supported. The coefficient of PM is –0.040 and is statistically significant at the 0.1 percent level. The point estimate is economically significant as well; evaluated at the mean, a one-standard-deviation increase in PM reduces the probability of exit by 7.2% percent, from 17.2% to 9.9%. As can be seen from the third column in Table 3, this change in the probability of exit is large compared to the other independent variables. The only variable with a comparable effect is industry growth. Moreover, the model performs significantly better than the model with only the control variables as regressors. The model \(\chi^2\) increases from 3,543.6 to 4,138.5 (significant at the 0.1 percent level) and the pseudo-\(R^2\) measure (Nagelkerke \(R^2\)) increases about 60 percent. Moreover, the fourth column in Table 3 shows how much explanatory power each variable adds to the model; PM adds more than any other variable, including such “classics” as industry growth, industry concentration, industry profitability and market share.
In total we interpret these findings as strong support for H1, and thereby (preliminary) support for the validity of the survivor principle in the context of diversification. We now move on to test whether the support for H1 is due to other factors than efficiency which would invalidate our interpretation of support for H1 as support for the survivor principle. H2 examines whether institutional isomorphism can drive the support for H1. As discussed above, the particular mechanism institutional theorists have suggested operated in our sample period implies that the support for H1 should be stronger in a sample of listed firms than in a sample of unlisted firms. Table 4 presents our findings when our sample is split into listed and unlisted firms.

Moving from Model 1 to Model 2 we can first conclude that the overall pattern we found in Table 3 holds for both listed and unlisted firms. For both subsamples in Table 4, adding PM increases the explanatory power of the model substantially, and the coefficients on PM are both negative and of similar magnitude as in Table 3. Actually, the coefficient is slightly larger for the subsample of unlisted firms, in contrast to what institutional isomorphism would predict. However, the standard error of the coefficient is larger for the unlisted firms, and as a result the ∆ Chi-square as we move from Model 1 to Model 2 is smaller for the unlisted firms. This difference in standard errors is more consistent with the isomorphism hypothesis (H2).

To further examine the role of institutional pressure we performed an additional test, relying on the argument that isomorphism should not only affect exit decisions, but equally—or possibly more so—entry decisions. As with exit decisions, listed firms should be exposed to stronger and earlier pressure to focus, making them less likely to make unrelated entry decisions. To examine this we sampled all the entry decisions made by diversified firms between 1983 and 1985 (replication for entries between 1981 and 1983 produced identical results). For each actual entry made by a diversifying firm we randomly drew four industries that the parent was not in or did not enter. We then regressed the log odds of entry on
PM and the appropriate control variables, which amounts to trying to predict which of the five potential target industries was the one actually entered. These regressions are reported in Table 5.

[Table 5 about here]

As Table 5 shows the coefficient on PM is 0.106 for unlisted firms and 0.074 for listed firms (controlling for initial firm size and diversity). This difference is the opposite of what one would expect if diversification behavior were shaped by institutional pressures operating through stock markets and institutional ownership. Note however, that here too the standard error of the coefficient is larger for the unlisted firms, so the estimate is less precise. Taken together, the results of these entry and exit regressions cast strong doubt on H2. In other words, support for H1 suggests that PM captures the efficiency of business combinations, not merely the quest for legitimacy.

H3 addresses the possibility that our findings reflect mutual forbearance rather than efficiency. To examine this we constructed a product term interacting industry concentration and PM. In Table 6 the regression without this product term is found in Model 2 and the regression with the product term included is in Model 3.

[Table 6 about here]

The interaction term is positive and significant, which at first glance suggests support for the mutual forbearance hypothesis. However, based on overall model performance, the interaction term does not appear substantively significant; the change in Nagelkerke $R^2$ is virtually zero, and an increase in $\chi^2$ of 13.82 from 4,138.53 is unimpressive. The significance of the coefficient and the change in $\chi^2$ is more a reflection of high statistical power than a substantively important finding. Hence while we cannot rule out mutual forbearance, it does not appear to be the main driver of our results. We also did several robustness

---

19 The independent variables are measured in the period preceeding the entry period. I.e. Industry growth is measured for 1881-83, the same for Industry profitability. Industry concentration is measured for the year 1983, as is Log parent size and parent diversity.
checks on this finding including centering the variables before forming the product term, and adding squared and cubed terms. Neither of these modifications performed better than the one reported in Table 6.

**CONCLUSIONS AND CAVEATS**

Our analysis shows that a survivor-based measure of relatedness is a good predictor of firms’ decisions to exit existing lines of business. This suggests that we can learn something about the efficiency of particular combinations of businesses by observing what diversified firms actually do. In other words, those businesses most frequently combined do on average represent more efficient combinations than those that are rarely combined. The comparative-efficiency version of the survivor principle thus withstands our attempt at falsification.

This finding should be interpreted carefully, however. First, our approach assumes that relatedness itself remained constant throughout our sample period. As our survivor-based measures are highly correlated between the 1981, 1983, and 1985 sample frames, this is probably a reasonable assumption. More generally, however, relatedness could vary over time, either exogenously (for example, through technological change), or endogenously, as new industries emerge and the entry of new firms changes the nature of relatedness between industries. Extensions to our work here would make explicit recognition of order and history; the relatedness between industries A and B might depend on whether firms already in industry A subsequently enter industry B, or vice versa.

Second, we cannot completely rule out alternative explanations for our results. As discussed above, isomorphism or some other form of herd behavior does not seem to be driving the observed pattern between survivor-based relatedness and exit. Mutual forbearance may play a role, but it appears to be a modest one. Third, our findings could reflect behaviors and patterns unique to the 1980s, a period often

Fourth, exit is only one measure of “efficient” behavior. Instead of focusing at the level of the industry pair and looking at changes in firm boundaries, one could also examine firm-level measures of overall relatedness—e.g., corporate-level relatedness scores as weighted averages of the neighbor relatedness scores for each of the firm’s operating units, which could then be used to explain profitability, firm growth, market value, and so on. These measures can also be compared with other firm-level relatedness measures such as sales- or asset-weighted Herfindahl indexes. Of course, our purpose here is not to contribute to the strategy literature on diversification, but to use diversification as a context for examining the survivor principle.

Fifth, though our findings support using the survivor principle in research on diversification, we cannot generalize this to say that relying on the survivor principle is OK in any setting or for studying any issue. On the other hand the fact that the survivor principle did survive testing in the context of corporate diversification should somewhat increase the a priori probability that it is valid in related areas—such as for example the study of vertical integration (where it is also commonly applied). Explicit examination of the survivor principle in such other contexts is, however, still necessary to judge its plausibility.

Finally, one may argue that even though it seems valid in a large economy such as the US, it may not be equally valid in smaller or less developed economies. And one may further argue that even though the findings supplied here indicate that it is valid for the private sector as a whole, there may be industries or settings within the private sector where it does not hold. Put differently, we have done more in terms of validating the survivor principle for use on large inter-industry samples than for smaller intra-industry samples.

---

20 See Matsusaka (1993), Hubbard and Palia (1999), and Klein (2001) for an alternative perspectives on the conglomerate period, however.
Obviously, much more work remains to be done in order to justify fully using the survivor principle in empirical studies. Nevertheless, in the present context, we conclude that the survivor principle did, indeed, survive diversification.

Our analysis has potentially important implications for organization and strategy research more generally. For example, while recent debates about the validity of transaction cost economics often focus on behavioral assumptions (in particular, opportunism and bounded rationality), the controversy is also largely empirical. Can organizational form really be interpreted as an attempt to economize on transaction costs, and does market competition really select for transaction cost efficiency (David and Han, 2004; Carter and Hodgson, 2006; Klein, 2005)? Key drivers of transaction cost efficiency such as asset specificity are notoriously difficult to parameterize and measure, however. An alternative is to look in large samples of firms for governance structures—patterns of vertical integration or long-term contracting, for instance—that persist over time under various conditions, and see if the reduced-form evidence on transaction-cost efficiency is consistent with the evidence from structural models that try to measure asset specificity directly and use it to explain performance.

It is also useful to understand how (mostly latent) characteristics such as inter-industry relatedness and transaction cost efficiency change in response to exogenous shocks, or across different institutional settings. The kind of analysis we demonstrate here could provide additional insight on the role of the institutional environment, and the effects of technology and policy shocks on business behavior. As noted above, much of the literature on the “round trip” of the US corporation from the 1960s to the 1990s assumes that relatedness itself remains roughly constant, and only agency problems, governance mechanisms, corporate control systems, and the like vary over time. But this assumption is not examined directly.

Survivor-based approaches may also be useful in helping us understand how the selection environment differs across contexts or over time. The recent financial crisis and subsequent downturn, and the regulatory changes that followed, have not only changed the “optimal” size and scope of financial and other firms, but changed the selection environment—making it tighter for construction and real-estate firms, for example, but (possibly) looser for large financial institutions, large manufacturers, and other big
players deemed “too big to fail.” This further strengthens our conviction that a better understanding of
selection and survival is critical for organization and strategy research.

REFERENCES

211–221.


Bain, J. S. 1956. *Barriers to new competition: Their character and consequences in manufacturing indus-


Benston, G. J. 1985. The Validity of Profits–Structure with Particular Reference to the FTC’s Line of


Bigelow, L., & Argyres, N. 2008. Transaction Costs, Industry Experience, and Make-or-Buy Decisions in
807.

Blaug, M. 1980. *The methodology of economics, or how economists explain*. Cambridge: Cambridge Uni-
versity Press.


### TABLE 1

Variable Definitions, Data Sources and Predicted Signs

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
<th>Data Source</th>
<th>Predicted Sign</th>
</tr>
</thead>
<tbody>
<tr>
<td>Industry growth</td>
<td>Sales growth in industry $i$ between 1981 and 1985</td>
<td>Trinet</td>
<td>$-$</td>
</tr>
<tr>
<td>Industry concentration</td>
<td>4-firm concentration ratio in industry $i$ (1981+1983)/2</td>
<td>Trinet</td>
<td>$-$</td>
</tr>
<tr>
<td>Industry profitability</td>
<td>$\Sigma$ industry return/$\Sigma$ industry assets, 1980–83 in industry $i$</td>
<td>Compustat</td>
<td>$-$</td>
</tr>
<tr>
<td>Market share</td>
<td>Market share in industry $i$ for parent $k$ in 1981</td>
<td>Trinet</td>
<td>$-$</td>
</tr>
<tr>
<td>Parent size</td>
<td>Ln (Total sales of parent $k$ in 1981)</td>
<td>Trinet</td>
<td>$-$</td>
</tr>
<tr>
<td>Parent diversity</td>
<td>Number of 4-digit industries parent $k$ participates in, in 1981</td>
<td>Trinet</td>
<td>$+$</td>
</tr>
<tr>
<td>$PM$</td>
<td>Weighted average survivor-based-relatedness of industry $i$ to the two closest neighboring industries in the portfolio of parent $k$</td>
<td>Trinet</td>
<td>$-$</td>
</tr>
</tbody>
</table>
**TABLE 2**
Means, Standard Deviations, and Correlation Coefficients for Independent Variables

<table>
<thead>
<tr>
<th>Variables</th>
<th>Mean</th>
<th>Standard deviation</th>
<th>Industry growth</th>
<th>Industry concentration</th>
<th>Industry profitability</th>
<th>Market share</th>
<th>Parent size</th>
<th>Parent diversity</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Industry variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 Industry growth</td>
<td>0.324</td>
<td>0.74</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Industry concentration</td>
<td>23.21</td>
<td>16.15</td>
<td>0.12**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Industry profitability</td>
<td>0.131</td>
<td>0.27</td>
<td>0.06** 0.07**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Firm variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 Market share</td>
<td>1.10</td>
<td>3.82</td>
<td>0.11** 0.36**</td>
<td>0.05**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 Parent size</td>
<td>7.85</td>
<td>1.90</td>
<td>0.05** 0.16**</td>
<td>0.01</td>
<td>0.18**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 Parent diversity</td>
<td>19.20</td>
<td>22.63</td>
<td>0.06** 0.10**</td>
<td>0.03**</td>
<td>0.10** 0.70**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SB-Relatedness variable</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 PM</td>
<td>22.64</td>
<td>15.51</td>
<td>0.08** 0.03**</td>
<td>0.04**</td>
<td>0.06** 0.12** 0.11**</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*a n = 39.403
* p < .05
** p < .01
**TABLE 3**

Logistic Regression Output on the Log Odds of Exit\(^{a}\)

<table>
<thead>
<tr>
<th></th>
<th>Model (1)</th>
<th>Model (2)</th>
<th>(\Delta p) in model 2(^{b})</th>
<th>(\Delta -2\text{LL}) per variable in model 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>−2.106 ***</td>
<td>−1.614 ***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.070)</td>
<td>(0.071)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Industry growth</td>
<td>−1.053 ***</td>
<td>−0.892 ***</td>
<td>−7.47%</td>
<td>1.309.95***</td>
</tr>
<tr>
<td></td>
<td>(0.027)</td>
<td>(0.027)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Industry concentration</td>
<td>0.007 ***</td>
<td>0.007 ***</td>
<td>1.61%</td>
<td>60.79***</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Industry profitability</td>
<td>−0.635 ***</td>
<td>−0.568 ***</td>
<td>−2.11%</td>
<td>54.24***</td>
</tr>
<tr>
<td></td>
<td>(0.082)</td>
<td>(0.082)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Market share</td>
<td>−0.096 ***</td>
<td>−0.076 ***</td>
<td>−3.75%</td>
<td>183.38***</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.007)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parent size (ln)</td>
<td>0.119 ***</td>
<td>0.149 ***</td>
<td>4.41%</td>
<td>234.98***</td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
<td>(0.010)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parent diversity</td>
<td>0.002 ***</td>
<td>0.004 ***</td>
<td>1.21%</td>
<td>23.20***</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PM</td>
<td>−0.040 ***</td>
<td></td>
<td>−7.20%</td>
<td>1,646.07***</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(-2\text{LL})</td>
<td>37,717.86</td>
<td>36,122.92</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model (\chi^2)</td>
<td>3,543.60***</td>
<td>4,138.53***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\Delta \chi^2) vs. model 1</td>
<td></td>
<td>1,595.94***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nagelkerke (R^2)</td>
<td>9.8%</td>
<td>15.6%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^{a}\) Logistic regressions of the log odds of exit on relatedness (PM) and industry and parent characteristics. \(n = 39,403\). Standard errors in parentheses.

\(^{b}\) Effect on the probability of exit of increasing each variable by 1 SD from its mean in model 2, while all other independent variables are held at their means.
TABLE 4

Logistic Regression Output on the Log Odds of Exit: Listed vs. Unlisted Firms

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th></th>
<th>Model 2</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unlisted</td>
<td>Listed</td>
<td>Unlisted</td>
<td>Listed</td>
</tr>
<tr>
<td>Constant</td>
<td>−2.565***</td>
<td>0.945***</td>
<td>−2.001***</td>
<td>−0.679***</td>
</tr>
<tr>
<td></td>
<td>(0.141)</td>
<td>(0.093)</td>
<td>(0.145)</td>
<td>(0.095)</td>
</tr>
<tr>
<td>Industry growth</td>
<td>−1.799***</td>
<td>−0.842***</td>
<td>−1.510***</td>
<td>−0.726***</td>
</tr>
<tr>
<td></td>
<td>(0.059)</td>
<td>(0.030)</td>
<td>(0.059)</td>
<td>(0.030)</td>
</tr>
<tr>
<td>Industry concentration</td>
<td>0.012***</td>
<td>0.004***</td>
<td>0.011***</td>
<td>0.004***</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.001)</td>
<td>(0.002)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Industry profitability</td>
<td>−0.480***</td>
<td>−0.964***</td>
<td>−0.410***</td>
<td>−0.822***</td>
</tr>
<tr>
<td></td>
<td>(0.136)</td>
<td>(0.122)</td>
<td>(0.138)</td>
<td>(0.121)</td>
</tr>
<tr>
<td>Market share</td>
<td>−0.085***</td>
<td>−0.092***</td>
<td>−0.069***</td>
<td>−0.073***</td>
</tr>
<tr>
<td></td>
<td>(0.020)</td>
<td>(0.008)</td>
<td>(0.018)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>Log (parent size)</td>
<td>0.142***</td>
<td>0.003</td>
<td>0.157***</td>
<td>0.053***</td>
</tr>
<tr>
<td></td>
<td>(0.022)</td>
<td>(0.012)</td>
<td>(0.023)</td>
<td>(0.012)</td>
</tr>
<tr>
<td>Parent diversity</td>
<td>−0.015***</td>
<td>0.005***</td>
<td>−0.003</td>
<td>0.005***</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.001)</td>
<td>(0.003)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>PM</td>
<td>−0.040***</td>
<td>−0.038***</td>
<td>0.002</td>
<td>(0.001)</td>
</tr>
</tbody>
</table>

\[−2 \text{ LL} \]

\[
\begin{align*}
\text{Model } \chi^2 & = 1,281.21*** \\
\Delta \chi^2 \text{ vs. model 1} & = 473.07*** \\
\text{Nagelkerke } R^2 & = 15.9\%
\end{align*}
\]

\[
\begin{align*}
\text{Model } \chi^2 & = 1,380.97*** \\
\Delta \chi^2 \text{ vs. model 1} & = 848.97*** \\
\text{Nagelkerke } R^2 & = 7.8\%
\end{align*}
\]

\[
\begin{align*}
\text{Model } \chi^2 & = 1,754.27*** \\
\Delta \chi^2 \text{ vs. model 1} & = 2,229.93*** \\
\text{Nagelkerke } R^2 & = 21.5\%
\end{align*}
\]

\[
\begin{align*}
\text{Model } \chi^2 & = 26,202.71 \\
\Delta \chi^2 \text{ vs. model 1} & = 848.97*** \\
\text{Nagelkerke } R^2 & = 12.5\%
\end{align*}
\]

\* \( p < 0.1 \)

\** \( p < 0.05 \)

\*** \( p < 0.01 \)

\( ^a \) Log odds of exit regressed on relatedness (PM) and industry and parent characteristics for listed and unlisted firms. Standard errors in parentheses. \( N = 13,867 \) (unlisted) and 25,536 (listed).
**TABLE 5**

Logistic Regression Output on the Log Odds of Entry: Listed vs. Unlisted Firms

<table>
<thead>
<tr>
<th></th>
<th><strong>Model 1</strong></th>
<th></th>
<th><strong>Model 2</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Unlisted</strong></td>
<td><strong>Listed</strong></td>
<td><strong>Unlisted</strong></td>
<td><strong>Listed</strong></td>
</tr>
<tr>
<td>Constant</td>
<td>–0.697 ***</td>
<td>–0.899 ***</td>
<td>–1.685 ***</td>
<td>–1.536 ***</td>
</tr>
<tr>
<td></td>
<td>(0.208)</td>
<td>(0.114)</td>
<td>(0.264)</td>
<td>0.133</td>
</tr>
<tr>
<td>Industry growth</td>
<td>0.158 ***</td>
<td>0.216 ***</td>
<td>0.011</td>
<td>0.022</td>
</tr>
<tr>
<td></td>
<td>(0.056)</td>
<td>(0.033)</td>
<td>(0.073)</td>
<td>0.039</td>
</tr>
<tr>
<td>Industry concentration</td>
<td>–0.028 ***</td>
<td>–0.020 ***</td>
<td>–0.027 ***</td>
<td>–0.019 ***</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.001)</td>
<td>(0.002)</td>
<td>0.001</td>
</tr>
<tr>
<td>Industry profitability</td>
<td>–0.004</td>
<td>–0.080 *</td>
<td>0.059</td>
<td>–0.039</td>
</tr>
<tr>
<td></td>
<td>(0.061)</td>
<td>(0.041)</td>
<td>(0.059)</td>
<td>0.039</td>
</tr>
<tr>
<td>Log (Parent size)</td>
<td>0.010</td>
<td>0.008</td>
<td>–0.030</td>
<td>–0.056 ***</td>
</tr>
<tr>
<td></td>
<td>(0.032)</td>
<td>(0.014)</td>
<td>(0.041)</td>
<td>0.016</td>
</tr>
<tr>
<td>Parent diversity</td>
<td>0.000</td>
<td>0.000</td>
<td>–0.048 ***</td>
<td>–0.012 ***</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.001)</td>
<td>(0.009)</td>
<td>0.001</td>
</tr>
<tr>
<td>PM</td>
<td></td>
<td></td>
<td>0.106 ***</td>
<td>0.074 ***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.004)</td>
<td>0.001</td>
</tr>
</tbody>
</table>

–2 LL                  | 4,904.40     | 19,027.63 | 3,441.25    | 15,001.578 |
Model $\chi^2$         | 245.93       | 588.61    | 1,709.08    | 4,614.67   |
$\Delta \chi^2$ vs. model 1 | 1,463.15 | 4,026.07 |
Nagelkerke $R^2$       | 7.4%         | 4.7%      | 44.7%       | 33.2%      |

* $p < 0.1$
** $p < 0.05$
*** $p < 0.01$

*a Log odds of entry regressed on relatedness (PM) and industry and parent characteristics for listed and unlisted firms. Standard errors in parentheses. $N = 5,123$ (unlisted) and 19,527 (listed)."
TABLE 6

Logistic Regression Output: Interaction Effect Between Industry Concentration and PM$^a$

<table>
<thead>
<tr>
<th></th>
<th>Model (2)</th>
<th>Model (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>$-1.614$ ***</td>
<td>$-1.511$ ***</td>
</tr>
<tr>
<td></td>
<td>$(0.071)$</td>
<td>$(0.076)$</td>
</tr>
<tr>
<td>Industry growth</td>
<td>$-0.892$ ***</td>
<td>$-0.894$ ***</td>
</tr>
<tr>
<td></td>
<td>$(0.027)$</td>
<td>$(0.027)$</td>
</tr>
<tr>
<td>Industry concentration</td>
<td>$0.007$ ***</td>
<td>$0.002$</td>
</tr>
<tr>
<td></td>
<td>$(0.001)$</td>
<td>$(0.001)$</td>
</tr>
<tr>
<td>Industry profitability</td>
<td>$-0.568$ ***</td>
<td>$-0.552$ ***</td>
</tr>
<tr>
<td></td>
<td>$(0.082)$</td>
<td>$(0.082)$</td>
</tr>
<tr>
<td>Market share</td>
<td>$-0.076$ ***</td>
<td>$-0.078$ ***</td>
</tr>
<tr>
<td></td>
<td>$(0.007)$</td>
<td>$(0.007)$</td>
</tr>
<tr>
<td>Parent size (ln)</td>
<td>$0.149$ ***</td>
<td>$0.150$ ***</td>
</tr>
<tr>
<td></td>
<td>$(0.010)$</td>
<td>$(0.010)$</td>
</tr>
<tr>
<td>Parent diversity</td>
<td>$0.004$ ***</td>
<td>$0.004$ ***</td>
</tr>
<tr>
<td></td>
<td>$(0.001)$</td>
<td>$(0.001)$</td>
</tr>
<tr>
<td>$PM$</td>
<td>$-0.040$</td>
<td>$-0.046$ ***</td>
</tr>
<tr>
<td></td>
<td>$(0.001)$</td>
<td>$(0.002)$</td>
</tr>
<tr>
<td>Concentration $^*$ $PM$</td>
<td>$0.000$ ***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$(0.000)$</td>
</tr>
</tbody>
</table>

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$-2$ Log likelihood</td>
<td>36,122.92</td>
<td>36,109.10</td>
</tr>
<tr>
<td>Model $\chi^2$</td>
<td>4,138.53***</td>
<td>4,152.35***</td>
</tr>
<tr>
<td>$\Delta \chi^2$ vs. model 2</td>
<td>13.82***</td>
<td></td>
</tr>
<tr>
<td>Nagelkerke $R^2$</td>
<td>15.6%</td>
<td>15.6%</td>
</tr>
</tbody>
</table>

$^a$ Log odds of exit regressed on relatedness ($PM$) and industry and parent characteristics. Model 3 includes an interaction term between industry concentration and $PM$. Standard errors in parentheses. $N = 39,403$