

**THE DETERMINANTS OF INDUSTRY CONCENTRATION:
TWO NEW EMPIRICAL REGULARITIES**

Lasse B. Lien*

Department of Strategy and Management
Norwegian School of Economics and Business Administration
Breiviksveien 40; N-5045 Bergen; Norway
Tel: +47 55 959726
Fax: +47 55 959780
lasse.lien@nhh.no

Nicolai J. Foss

Department of Strategy and Management
Norwegian School of Economics and Business Administration
Breiviksveien 40; N-5045 Bergen; Norway
Tel: +45 3815 2562
Fax: +45 3815 2500
njf.smg@cbs.dk

and

Center for Strategic Management and Globalization
Copenhagen Business School
Porcelainshaven 24, 1st fl.
2000 Frederiksberg; Denmark

Forthcoming in *Managerial and Decision Economics*

* Corresponding author

THE DETERMINANTS OF INDUSTRY CONCENTRATION:

TWO NEW EMPIRICAL REGULARITIES

KEYWORDS: Industry concentration, entry, relatedness, diversification, stylized fact.

ABSTRACT

This paper reports two new empirical regularities relating to industry concentration. First, concentration levels closely correlate in related industries. Second, the correlation is moderated by the degree of relatedness between the industries. These regularities are derived from the Trinet database, using a survivor-based measure of relatedness. We argue that these previously overlooked relations may be explained in terms of 1) “spillover effects” between industries and 2) lifecycle factors.

INTRODUCTION

Industries have neighbors, that is, other industries that are closely related to the focal industry. The relation between the structural characteristics of such neighbor or related industries, as well as the implications of this for firm behavior, has not been offered much attention in industrial economics and strategic management.¹ For example, in investigating entry behavior, industrial economists have generally abstracted from how the industry affiliation of the potential entrant may influence its competitive behavior and chances of successful entry.

In this paper we report two hitherto neglected empirical regularities concerning the relations between the structural characteristics of industries, specifically concentration levels. First, we find that concentration correlates on a surprisingly high level ($> .20$) across related industries, and, secondly, that the strength of the correlation depends on the relatedness of the industries.

The discovery of empirical regularities emerges from a context of discovery (Popper, 1934). In our case, the relevant context is the following two stylized facts (Helfat, 2007). First, an important finding from the strategy and industrial organization literatures is that firms' diversification behaviors display a systematic preference for closely related industries (Lemelin, 1982; Chatterjee and Wernerfelt, 1991; Montgomery and Hariharan, 1991; Silverman, 1999).² Second, industry affiliation matters to the success of entry, and therefore also for the evolution of the structure of the target industry. Thus, another important finding from research in strategy and

¹ A partial exception is constituted by work on multipoint competition (Edwards, 1955; Bulow et al., 1985; Karnani and Wernerfelt, 1985; Gimeno, 1999; Gimeno & Woo, 1999).

² Note that the findings regarding relatedness and the direction of diversification are much more robust and stable than the findings regarding the link between relatedness and corporate performance (e.g., Montgomery & Hariharan, 1991; Silverman, 1999).

industrial organization is that diversifying entrants are more likely to survive and grow than newborn entrants (i.e., entrants without industry affiliation) (Baldwin, 1995; Dunne, Roberts, & Samuelson, 1989; Siegfried & Evans, 1994; Geroski, 1995; Sharma & Kesner, 1996).

Combining these two stylized facts implies that diversifying entrants pose a bigger threat for incumbents in terms of loss of market share and intensified rivalry than newborn entrants (*ceteris paribus*). This suggests that relations between neighboring industries matter to entry behavior and to *ex post* competition. This may show up in the data as statistical relations among the key structural characteristics of neighboring industries. A natural place to start empirical inquiry into such issues is to investigate industry concentration, because of its status as the perhaps key structural industry characteristic in the strategic management and industrial organization literatures. Such reasoning led to the finding of the two regularities reported in this paper.

The regularities are derived from the AGSM/Trinet database which is a database covering the entire US economy in the 1980s, offering biannual data on the establishment level. We find that a large portion of the variance in industry concentration can in fact be explained by the concentration that characterizes related industries and the degree of relatedness between the focal industry and its neighbors. We briefly develop a set of possible explanations of these regularities; however, the data set does not allow us to directly test such explanations.

In sum, the contribution of this paper is to report two new regularities between industry concentration levels, and offer possible theoretical explanations of the regularities. To our knowledge, this has not been done previously in the industrial economics or strategic management literatures.

A SURVIVOR-BASED APPROACH

Empirically examining correlations between structural characteristics of industries raises huge problems that have to do with separating the relevant from the irrelevant industries. In other words, which industries are likely to influence each other in terms of structural characteristics? As suggested, existing empirical findings provide clues here, particularly the finding that firms prefer to diversify into related industries (e.g., Montgomery & Hariharan, 1991). This leads to the prediction that industries differ with respect to the impact on concentration from diversifying entry and that this difference may be explained in terms of relatedness. The problem is, however, that the conventional procedures for measuring the relatedness between industries are too coarse: Using distances within the SIC-hierarchy leaves one with only three levels of distance, for example, different 4-digit but identical 3-digit industry, different 3-digit but identical 2-digit industry, different 2-digit industries. This would not allow us to identify which industries are closest, and certainly not variation in the distance to those designated as the closest neighbors (all industries would have a large amount of equally close neighbors). In other words, a different and preferably continuous measure of relatedness is called for.

A *survivor-based measure* of relatedness is a candidate for such a measure. This measure means that relatedness is determined by how often a given pair of industries is actually combined within a firm, compared to what one would expect if diversification patterns were random (Lien & Klein, 2004; Teece et al., 1994). A pair of industries are related to the degree that this

difference is positive and unrelated to the degree that it is negative). The measure thus produces a continuous survivor-based metric of the relatedness between all industries in the economy.

The term “survivor-based” refers to the fact that the measure is based on observing the outcome of decisions by those with superior knowledge, namely local decisionmakers, and the screening of those decisions by the competitive process. In other words, the measure picks up whatever matters for decision makers as well as the viability of a combination of industries. This means that we are agnostic about what drives relatedness for a particular combination of industries, which constitutes both a strength and a weakness. The strength is that it allows the causes of relatedness to vary across situations. It does not involve singling out any particular class of assets, for example, customer related assets, technological assets or production assets (or some combination of these) (Farjoun, 1994; Montgomery & Hariharan, 1991; Silverman, 1999) as *the* universal source of economies of scope or complementarity. Rather, this may vary from case to case. The main weakness is, of course, that the measure may be noisy due to contamination from motives and mechanisms unrelated to relatedness, such as herd-behavior or empire-building. However, attempts to validate the survivor based measure have shown it to consistently and significantly outperform SIC-based measures (references withheld).

Calculating a Survivor-based Measure

Our technical approach is based on a procedure developed by Teece et al. (1994). Let K be the universe of diversified firms, and I be the universe of industries these firms are active in. Let $C_{ik} = 1$ if firm k is active in industry i . The number of diversified firms present in industry i is then $n_i = \sum_k C_{ik}$. Let J_{ij} be the number of diversified firms participating in both industries i and j , such that $J_{ij} = \sum_k C_{ik}C_{jk}$. J_{ij} will be larger if industries i and j are related, but is also increasing 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 expected number of combinations if diversification patterns were random.

To calculate the expected number of combinations we consider a hyper-geometric 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:

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

The mean and variance of X_{ij} are:

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

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

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

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

Given this basic measure of the relatedness between a pair of businesses, it is possible to rank, for any industry, how close all other industries are, and it is possible to calculate aggregate measures of how close, say, the four closest industries are to a given industry.

To calculate SR_{ij} we used the AGSM/Trinet Large Establishment Database (Trinet). The Trinet database covers the 1980s, containing biannual records of all US establishments with more than 20 employees, including variables such as 4-digit SIC code, corporate ownership, and sales.³ The primary measure of SR_{ij} is calculated from the Trinet files of 1981, 1983, 1985 and 1987 using all recorded firms active in two or more 4-digit SIC codes as a basis.

Aggregation of Variables

To understand how industry concentration in one industry is influenced by concentration in neighboring industries, it is necessary to derive summary measures that aggregate variables over all industries close enough to matter. It is not sufficient to merely focus on the closest neighbor, since, for example, if the closest industry is concentrated, there can be a second and third industry almost equally close that are fragmented. A decision concerning how many industries we should sum over is required. In order to determine what a reasonable cut-off point might be, we resort to the following simple procedure: We use the concentration ratio, measured as the C4 ratio of the focal industry as the dependent variable, and then introduce the concentration ratio in the closest-, second closest-, third closest- (and so on) industry as independent variables. Table 1 provides the regression output for 1981. As can be inferred from the Table, the concentration ratios of the four closest industries are all significant predictors of concentration in the focal

³ 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 0,893 which indicates that the sales data in Trinet are of acceptable quality.

industry, while the fifth closest industry is not. We therefore use the four closest industries as the basis for computing aggregated variables.

Insert Table 1 here

We measure the concentration of neighbouring industries in two different ways. One measure is the sum of the C4 ratios of the four closest industries (*C4Neighbors*). The other, complementary, measure is the number of potential diversifying entrants from the closest neighboring industries. We calculate this as the number of firms in the neighboring industry, minus the number of firms already combining the two industries. Summing this over the four closest industries we obtain a measure of the size of the pool of diversifying entrants from these industries (*Entrypool*). The reason for including both is that the C4 measure alone is insensitive to the structure of the smaller firms in an industry. For example, a C4 ratio of 0.80 does not reveal whether the remaining 20% is divided between 2 or 50 firms. For our purposes this is possibly important, since these smaller firms represent potential diversifying entrants.

Above, we described our procedure for calculating the fundamental measure of SR_{ij} . The variable *Relatedness* is the simple sum of SR_{ij} for the four closest industries. The purpose of this variable is twofold. First, it serves to identify neighbor industries. Second, it is used to examine the moderation effect of relatedness on the strength of the correlation between concentration levels across industries.

The interaction variable is measured in two different ways. The first one multiplies *C4Neighbors* and *Relatedness*, and is termed *C4XRelatedness*. The second uses the size of the

entry pool instead of the C4 ratio and is created by multiplying *Entrypool* and *Relatedness*. This term, *PoolXRelatedness*, should have the opposite sign of *C4XRelatedness*.

Finally, the dependent variable (*Industry concentration*) is the C4-ratio of the focal industry.

We use a sample from the Trinet database to examine the relationship between these variables. The sample is essentially population data for all 4-digit SIC industries in 1981, 1983, 1985 and 1987. The only industries excluded are nonprofit industries, and industries that do not have relationships to at least four other industries. This means that for example industries that have no diversified firms are excluded.⁴ This produces a sample size of 855 industries in 1981, 845 industries in 1983, 847 industries in 1985, and 837 industries in 1987.

We do not report the mean, standard deviation and correlations for all years, but Table 2 provides data for 1981 as a representative example.

Insert Table 2 here

MODELS AND RESULTS

To examine our data for empirical regularities we apply OLS regression. The first set of regressions are provided in Model 1 and Model 2 of Table 3. These regressions are designed to examine the direct (unmoderated) effect of concentration in neighboring industries on a focal industry. In Model 1 we only include the variable *Entrypool* and a constant term. It seems

⁴ Note that the presence of one diversified firm is sufficient for an industry to be included, provided that this one firm is also active in four other industries.

reasonable to expect the coefficient on *Entrypool* to be negatively signed, since a large pool of potential diversifying entrants is more likely to increase concentration than reduce it. As the results in Model 1 shows, our data confirm this expectation. In all four years the coefficient is negative and significant. In terms of effect size, the standardized coefficients show that one standard deviation increase in *Entrypool* leads to between 0,21 and 0,27 standard deviations reduced concentration. Model 2 adds the other concentration measure *C4Neighbors* to the model. If *Entrypool* is negatively signed, one would expect would *C4Neighbors* to be positively signed. As model 2 shows the coefficient on *C4Neighbors* is indeed positively signed and significant in all years. In terms of effect size, the standardized coefficients show that one standard deviation increase in *C4Neighbors* leads to between 0,32 and 0,43 standard deviations increased concentration. Note also that the coefficients on *Entrypool* are reduced in size, but remain significant in all years except 1987. Apparently these two measures are somewhat complementary, but they obviously overlap significantly. Among the two, *C4Neighbors* is clearly the better measure in terms of explaining concentration. The adjusted R^2 for model 2 ranges between 0,13 and 0,19 indicating that the structure of neighboring industries is a significant determinant of industry structure.

Insert Table 3 here

Table 4 examines how the degree of relatedness moderates the correlation between concentration levels. Table 4 contains two models. Model 3 adds the interaction term *PoolXRelatedness* to model 1, and Model 4 adds the interaction term *C4XRelatedness* to model

2. Generally, it seems reasonable to expect relatedness to strengthen the effects in Model 1 and Model 2. I.e., if increasing the pool of potential entrants from related industries reduces concentration, one would expect the reduction to be stronger the shorter the distance such entrants have to “jump”. In terms of model 3 this would imply a negative coefficient on the interaction variable *PoolXRelatedness*. The data support this. The coefficient is indeed negative and highly significant in all four years. It is also notable that the model with the interaction term (Model 3) performs significantly better than the model without it (Model 1) in all four years ($p < 0,001$). For Model 4 the same logic implies a positive coefficient on the interaction variable *C4XRelatedness*. Again, we find that the variable is signed as expected and significant all four years, and also that the model with the interaction term (Model 4) performs significantly better than the model without it (Model 2). There seems to be a The best model in terms of adjusted R^2 is model 4, ranging from 0,14 to 0,21.

So far we have found systematic evidence that concentration correlates quite strongly across related industries, and that this correlation is moderated by the degree of relatedness. We have not offered explanations for why this may be the case. We now move on to discuss possible explanations for these empirical regularities.

Insert Table 4 here

DISCUSSION

The structural characteristics of industries have long been central in strategic management and industrial organization research. Although the research emphasis in industrial organization

research has changed towards modeling structural characteristics as endogenous outcomes of strategic interaction (rather than taking them as exogenous variables), structural characteristics still play a key role as constraints on the equilibria that can be reached in a game. Similarly, strategic management research continues to make use of variables that capture structural effects as key independent variables in theoretical and empirical research (e.g., Sharma and Kesner, 1996; Gimeno, 1999). Arguably, *the* key structural variable remains that of industry concentration. However, in decades of research that has investigated the determinants and effects of concentration, no research has (to our knowledge) examined how concentration in one industry may influence concentration in another industry and what may moderate this. This is exactly where we contribute by reporting the above regularities.

Why Concentration Levels Correlate

A number of possible theoretical explanations exist that may account for the regularities reported here. The limitations of the data set mean that we cannot test these. However, we briefly outline these explanations in order to stimulate further research.

Consider the first reported regularity, namely that the level of concentration in an industry correlates positively with the level of concentration in its closest neighboring industries. There are three mechanisms that may (jointly, pair wise, or in isolation) establish such a correlation.

First, if the industries close to the focal industry are concentrated, there is a smaller pool of potential diversifying entrants (*ceteris paribus*). In other words, the threat of *direct* entry is smaller. This weakens an important mechanism that may otherwise contribute to reducing concentration. Second, concentrated neighboring industries are themselves likely to be difficult to enter. This reduces the number of entrants that can enter the focal industry *indirectly*, that is,

using neighboring industries as stepping stones.⁵ Third, high levels of economies of scope between neighboring industries can create an entry barrier that is *shared* between the industries, facilitating concentration in both the focal industry and its neighboring industries.

These effects of these mechanisms may be reinforced by industry life cycle effects. Work on industry life cycles (e.g., Klepper & Graddy, 1990) typically stress that nascent industries tend to be competitive and have relatively low entry barriers. However, industry evolution which is driven by technological development tends to be accompanied by a shake-out that increases concentration. To the extent that related industries make use of similar technology, they are likely to be at the same stages of their life-cycles, and are therefore likely to have similar same levels of concentration. While this argument would conceivably apply to entirely specialized firms in separate industries (thus offsetting the above mechanisms), we think it is more likely to reinforce the above mechanisms. Thus, industries that are related in terms of technology are also likely to exhibit economies of scope and resource complementarities (Teece, 1980; Teece et al., 1994). Such economies are often seen as important antecedents of diversification. Consequently, diversification that stems from similar technology across industries is likely to be captured by our survivor-based measure of relatedness.

The Moderating Effect of Relatedness

Consider also the second regularity, namely that the correlation between the levels of concentration in neighboring industries is positively moderated by their degree of relatedness. There are (at least) three mechanisms that may produce this regularity. First, relatedness impacts the entry decision (Penrose, 1959): The less related a potential target industry is, the less

⁵ Analogous to the way in which firms may use strategic groups within industries as stepping stones to the strategic group they ultimately prefer to stay within (Caves & Porter, 1977).

attractive it is for a potential entrant. Second, entry barriers based on economies of scope may play a role: The less related an industry is, the less the relevant economies of scope, and the less important it is to be active in both to be competitive in either. Third, again a life cycle argument may be relevant: If industries are related they are also likely to make use of similar technology, to be at a similar point of their evolution, and therefore to have similar levels of concentration.

Relevance

The findings and arguments in this paper contribute to the empirical industrial organization literature. Moreover, our arguments and findings hold implications for the strategic management literature, particularly the diversification literature. For example, consider the debate on diversification and performance (beginning with Rumelt, 1974; Bettis, 1981; Christensen & Montgomery, 1981; Rumelt, 1982). In their critique of Rumelt (1974) Christensen and Montgomery (1981) argued that the relatedness/performance link in (an updated version of) Rumelt's sample was strongly influenced by industry characteristics: Controlling for such characteristics largely eliminated Rumelt's (1974) finding of a positive relatedness/performance link. However, this line of argument takes industry characteristics as exogenously given, whereas we have suggested that one such characteristic (i.e., concentration) is endogenous to diversification decisions. This paper thus suggests that using industry concentration (and possibly also other industry characteristics) as a control in relatedness/performance studies is a possible source of bias, since the findings reported here are consistent with the notion that concentration is to a considerable extent an outcome of relatedness and diversification decisions.

Although industry concentration may be partly endogenous to firm actions, it is still a variable that is highly relevant to practicing strategists. First, these ideas allow strategists to

include variables that are relevant to assessing entry threats, namely the degrees of concentration of neighboring industries and the role that relatedness plays in influencing this threat. Second, strategists can use the ideas in this paper to structure thinking about how entry threats evolve as a function of changing degrees of concentration and relatedness. For example, a technological change that works to fragment a neighboring industry will tend to reduce concentration levels in the focal industry. How much, depends not only on how much the neighboring industry is fragmented, but also on how closely related the two industries are. A second example is a technological change that increases the relatedness to a neighboring industry. The effect of such a change will depend on the structure of this industry. If the industry in question is sufficiently fragmented, it will reduce concentration in the focal industry, while if it is sufficiently concentrated the opposite will occur.

Future Work

Future work on these issues may proceed in a number of directions. First, more empirical work on different data is needed to establish our findings as true “stylized facts,” that is, regularities that appear in so many contexts that they may be considered empirical truths (Helfat, 2007: 187). Thus, we have only examined US data for the 1980s.

Second, the theoretical mechanisms we have proposed as candidates for explaining the regularities need to be subject to empirical testing. This requires more fine-grained data than we have had access to.

Third, the theoretical mechanisms need to be formally modeled in a multi-period setting. For example, we have posited that diversifying entrants are the most dangerous species of entrants in terms of driving down concentration. While this may be true in an initial stage of a

multi-period game, concentration may increase in later stages. Concentration may be high in an industry because incumbents are highly efficient (Demsetz, 1973); when such firms enter a neighboring industry, they initially drive down concentration, but may later use their strengths to significantly take away market share from weaker incumbents of the industry, effectively increasing concentration. A further issue that needs to be theoretically as well empirically examined concerns the mode of entry, that is, do entrants enter by means of mergers and acquisitions or *de novo* entry. Consistent with the industrial economics literature and because of the constraints imposed by the dataset, we have not treated acquisitions as entry in this paper; however, it is obvious that for strategists considering an entry decision, the choice of entry mode is a pressing one. From the theory perspective, entry mode choices may affect, for example, the level of sunk cost commitment and therefore the form that *ex post* competition takes, and, in turn, the evolution of industry concentration.

REFERENCES

- Baldwin J. R. 1995. *The dynamics of industrial competition*. Cambridge University Press: Cambridge.
- Bettis, R.A. 1981. Performance differences in related and unrelated diversified firms. *Strategic Management Journal*, **2**: 379-393.
- Caves, R. & Porter, M.E. 1977. From entry barriers to mobility barriers: Conjectural decisions and contrived deterrence to new competition. *Quarterly Journal of Economics*: **91**:241-262.
- Chatterjee, S., & Wernerfelt, B. 1991. The link between resources and type of diversification: Theory and evidence. *Strategic Management Journal*, **12**: 33-48.
- Christensen, H.K. & Montgomery, C. 1981. Corporate economic performance: Diversification

- strategy versus market structure. *Strategic Management Journal*, **2**: 327-343.
- Davis, P.S., Robinson, R.B., Pearce, J.A., & Park, S.H. 1992. Business unit relatedness and performance: A look at the pulp and paper industry. *Strategic Management Journal*, **13**: 349-361.
- Demsetz, H. 1973. Industry structure, market rivalry, and public policy. *Journal of Law and Economics*, **16**: 1-9.
- Dunne T., Roberts M.J., & Samuelson L. 1989. The growth and failure of United States manufacturing plants. *Quarterly Journal of Economics* **104**(4): 671-698
- Farjoun, M. 1994. Beyond industry boundaries: Human expertise, diversification and resource-related industry groups. *Organization Science*, 5(2): 185–199.
- Geroski P.A. 1995. What do we know about entry? *International Journal of Industrial Organization*, **13**: 421-440.
- Gimeno, J. 1999. Reciprocal threats in multi-market competition: Staking out spheres of influence in the US airline industry. *Strategic Management Journal*, **20**: 101-128.
- Gimeno, J. & Woo, C.Y. 1999. Multimarket contact, economies of scope, and firm performance. *Academy of Management Journal*, **43**: 239-259.
- Helfat, C. 2007. Stylized facts, empirical research and theory development in management. *Strategic Organization*, **5**: 185-192.
- Jaccard, J. 2001. *Interaction effects in logistic regression*. London: Sage Publications.
- Karnani, A. & Wernerfelt, B. 1985. Multiple point competition. *Strategic Management Journal*, **6**: 87-96.
- Klepper, S. & Graddy, E. 1990. The evolution of new industries and the determinants of market structure. *RAND Journal of Economics*, **21**: 27-44.
- Lemelin, A. 1982. Relatedness in the patterns of inter-industry diversification. *Review of Economics and Statistics* 646-657.

- Lien, L. B. & Klein, P. 2004. Yet another way of measuring relatedness - This one: Let competition do it! *Best Paper Proceedings of the Annual Meeting of the Academy of Management*, New Orleans, USA.
- Markides, C. C. & Williamson, P. J. 1996. Corporate diversification and organizational structure: A resource-based view. *Academy of Management Journal*, **39**: 340–367.
- Montgomery, C.A. 1985. Product-market diversification and market power, *Academy of Management Journal*, **28**: 789-798.
- Montgomery, C. A., & Hariharan, S. 1991. Diversified expansion by large established firms. *Journal of Economic Behavior & Organization*, **15**: 71–89.
- Penrose, E.T. 1959. *The theory of the growth of the firm*. Oxford: Oxford University Press.
- Popper, K.R. 1934. *The logic of scientific discovery*. New York: Routledge.
- Robins, J., & Wiersema, M. F. 1995. A resource-based approach to the multi-business 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**(1): 39–59.
- Rumelt, R.P. 1974. *Strategy, structure, and economic performance*. Boston: Harvard University Press.
- Rumelt, R.P. 1982. Diversification strategy and profitability. *Strategic Management Journal*, **3**: 359-369.
- Sharma, A. 1998. Mode of entry and *ex post* performance. *Strategic Management Journal* **19**: 879-900.
- Sharma, A. & Kesner, I.F. 1996. Diversifying entry: Some *ex ante* explanations for postentry survival and growth, *Academy of Management Journal*, **39**: 635-677.
- Sigfried, J.J. & Evans, L.B. 1994. Empirical studies of entry and exit: A survey of the evidence.

Review of Industrial Organization, **9**: 121-155.

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**(8): 1109–1124.

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., Dosi, G., Rumelt, R.P., & Winter, S.G. 1994. Understanding corporate coherence. *Journal of Economic Behavior and Organization*, **23**: 1-30.

Table 1: OLS Regression of concentration (C4) within an industry by concentration (C4) in neighboring industries, 1981 (N = 855)

	Unstandardized Coefficient (S.E)	Standardized Coefficient	t-statistics
Constant	11,371*** (2,029)		5,605
C4 closest	0,210*** (0,032)	0,217	6,607
C4 2nd closest	0,155*** (0,031)	0,164	4,998
C4 3rd closest	0,163*** (0,033)	0,159	4,992
C4 4th closest	0,152*** (0,032)	0,152	4,752
C4 5th closest	-0,002 (0,031)	-0,002	0,073

*p < 0.05, **p < 0.01, ***p < 0.001

Table 2: Means, standard deviations and correlation coefficients of independent variables, 1981 (N = 855)

	Mean	SD	C4Neighbors	Entrypool	Relatedness	C4XRelatedness
C4Neighbors	139,40	56,00	1			
Entrypool	295,32	338,81	-0,347***	1		
Relatedness	82,78	39,54	0,150***	-0,244***	1	
C4XRelatedness	11591,9	7960,4	0,665***	-0,322***	0,760***	1
PoolXRelatedness	22140,1	24716,2	-0,318***	0,609***	0,234***	-0,088**

*p < 0.05, **p < 0.01, ***p < 0.001

Table 3: OLS Regression of concentration (C4) within an industry by concentration in neighboring industries (N = 855,845, 847, 837)

			1981	1983	1985	1987 ⁶
Model 1	Constant		39,54*** (1,028)	39,36*** (1,09)	41,76*** (1,15)	0,59*** (0,1)
	Entrypool	Unstandardized coeff. (S.E.)	-0,015*** (0,002)	-0,015*** (0,002)	-0,020*** (0,003)	-0,000*** (0,000)
		Standardized Coefficient	-0,224***	-0,212***	-0,236***	-0,274***
	Model 1 Adjusted R²		0,049	0,044	0,054	0,074
Model 2	Constant		14,88*** (2,38)	17,32*** (2,47)	14,63*** (2,03)	0,19*** (0,04)
	Entrypool	Unstandardized coeff. (S.E.)	-0,006** (0,002)	-0,006* (0,002)	-0,007* (0,003)	0,000 (0,000)
		Standardized Coefficient	-0,089**	-0,080*	-0,082*	0,001
	C4Neighbors	Unstandardized coeff. (S.E.)	0,158*** (0,014)	0,141*** (0,014)	0,133*** (0,015)	0,163*** (0,015)
		Standardized Coefficient	0,378***	0,339***	0,318***	0,435***
	Model 2 Adjusted R²		0,173	0,141	0,131	0,187

*p < 0.05, **p < 0.01, ***p < 0.001

⁶ Note that the SIC-system was revised in 1987 which affected about 30% of the relevant industries. Some industry definitions were widened, and some narrowed. This means caution is needed when comparing 1987 figures to the other years.

Table 4: OLS Regression of concentration (C4) within an industry by concentration-, relatedness- and interaction between concentration and relatedness in neighboring industries (N = 855,845, 847, 837)

			1981	1983	1985	1987⁷
Model 3	Constant	Unstandardized Coefficient (S.E.)	31,88*** (2,20)	32,75*** (2,44)	35,81*** (2,56)	0,61*** (0,026)
	Entrypool	Unstandardized coefficient (S.E.)	0,002 (0,004)	0,004 (0,005)	-0,003 (0,006)	0,000 (0,000)
		Standardized Coefficient	0,023	0,049	-0,033	0,008
	Relatedness	Unstandardized coefficient (S.E.)	0,102*** (0,024)	0,080*** (0,025)	0,072** (0,027)	0,000 (0,000)
		Standardized coefficient	0,174***	0,145***	0,127**	0,001
	PoolXRelatednes	Unstandardized coefficient (S.E.)	-0,000*** (0,000)	-0,000*** (0,000)	-0,000*** (0,000)	-0,000*** (0,000)
		Standardized coefficient	-0,280***	-0,300***	-0,241***	-0,387***
	Model 3 Adjusted R²		0,073	0,059	0,065	0,142

*p < 0.05, **p < 0.01, ***p < 0.001

⁷ Note that the SIC-system was revised in 1987 which affected about 30% of the relevant industries. . Some industry definitions were widened, and some narrowed. This means caution is needed when comparing 1987 figures to the other years.

Table 4 (cont.)

			1981	1983	1985	1987
Model 4	Constant	Unstandardized Coefficient (S.E.)	27,02*** (5,01)	30,18*** (4,87)	25,47*** (5,31)	0,492*** (0,08)
	Entrypool	Unstandardized coefficient (S.E.)	-0,005* (0,002)	-0,005* (0,002)	-0,006 (0,003)	-0,000 (0,000)
		Standardized coefficient	-0,081*	-0,074*	-0,069	-0,067
	C4Neighbors	Unstandardized coefficient (S.E.)	0,042 (0,032)	0,031 (0,031)	0,069* (0,032)	0,050 (0,032)
		Standardized coefficient	0,101	0,074	0,165*	0,132
	Relatedness	Unstandardized coefficient (S.E.)	-0,141* * (0,050)	-0,145** (0,046)	-0,086 (0,052)	-0,003*** (0,001)
		Standardized coefficient	-0,239**	-0,263**	-0,152	-0,495***
	C4XRelatednes	Unstandardized coefficient (S.E.)	0,001*** (0,000)	0,001*** (0,000)	0,001* (0,000)	0,003*** (0,001)
		Standardized coefficient	0,451***	0,427***	0,267*	0,447***
	Model 3			0,195	0,157	0,137
Adjusted R²						

*p < 0.05, **p < 0.01, ***p < 0.001