A Logical Approach to Competition in Industries

Ivar Vermeulen

A Logical Approach to Competition in Industries

ILLC Dissertation Series DS-2002-06

INSTITUTE FOR LOGIC, LANGUAGE AND COMPUTATION

For further information about ILLC-publications, please contact

Institute for Logic, Language and Computation
Universiteit van Amsterdam
Plantage Muidergracht 24
1018 TV Amsterdam
phone: +31-20-525 6051
fax: +31-20-525 5206
e-mail: illc@wins.uva.nl
homepage: http://www.illc.uva.nl/

A Logical Approach to Competition in Industries

Academisch Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof.mr. P.F. van der Heijden
ten overstaan van een door het college voor
promoties ingestelde commissie, in het openbaar
te verdedigen in de Aula der Universiteit
op dinsdag 6 februari 2003, te 10.00 uur

door

Ivar Erik Vermeulen

geboren te Amsterdam.

Promotiecommissie:

Promotor: Prof.dr. W.E. Saris
Co-promotor: Dr. J.M.F. Masuch
Faculteit der Maatschappij- en Gedragswetenschappen

Overige leden:
Prof.dr. J.F.A.K. van Benthem
Prof.dr. C. Boone
Prof.dr. K.M. Carley
Prof.dr. R.J. Mokken
Prof.dr. J. Visser

Cover design by Studio Dick de Vries
Printed and bound by PrintPartners Ipskamp, Enschede

ISBN: 90–5776–096–7

Voor Erik en Jacobine

Contents

Acknowledgments

I would like to thank my supervisor, Willem Saris, for cheerfully accepting me, together with my odd range of subjects, as a Ph.D student, and for his thoughtful comments and suggestions. My daily supervisor, Michael Masuch, I would like to thank for his confidence, his day-to-day support, and his sparse, but ever so critical interventions in my research.
Lots of thanks to Jeroen Bruggeman, my collaborator on many occasions, and co-author of two chapters of this dissertation. Without his supervision, I would still be working on My First Paper.
Thanks to Jaap Kamps for helping me out on almost a daily basis, for criticizing my various stupid ideas, and for enhancing the quality of my manuscripts. Thanks also to Rob Mokken, for continually stressing the importance of details, and for checking numerous ones in this manuscript.
Thanks to all my colleagues at ILLC/LIT, for making it such a fun place to work.

Thanks to my dear old friends for patiently listening to my ramblings. Thanks to the guys in the band and in the various football squads for reminding me of what is important in life.
Thanks to my family, and to my "in-laws," for their unfounded yet persistent confidence in me and in my abilities.

Thanks, finally, most of all, to Sanne—not so much for her contribution to this thesis, as for her contribution to my life and happiness.

Amsterdam — Ivar Vermeulen
December, 2002

Chapter 1

Introduction

Many texts that present social science theory do not meet standards of being unambiguous and logical (Sutton and Staw 1995). These texts leave the burden of understanding and disambiguating theory, and of determining the theory's meaning and logic, on the readers. Texts that lack clarity do not fully exploit the ideas they aim to present. As a consequence, the ideas may be misunderstood, or fall into oblivion.

To increase a theory's rigor and precision, and to raise the discussion to a new level, representing a theory in formal logic is an approved method (Suppes 1968; Kyburg and E. 1968; Péli et al. 1994; Hannan 1997; Kamps and Pólos 1999). Logical formalization does not only facilitate evaluating the soundness of a theory, it also forces to extract the theory's "core," to determine the domain to which the theory applies, to disambiguate its concepts, and to specify the implicative structure of its argument.

Formalizing a text containing theory in logic is a non-trivial affair. If representing natural language in formal logic were trivial, logic would not be needed at all; natural language would provide sufficient footing for evaluating arguments. Although formalization is non-trivial, approaching it in a systematic manner may make it easier. In the second chapter of this thesis, we present a systematic approach to (logical) formalization of social science theories.

In the third chapter, we show the use of our approach by formalizing a part of organizational ecology, *i.e.*, resource partitioning theory (Carroll 1985), in first-order logic. We chose resource partitioning theory as the subject of a formalization attempt partly, because various other parts of organizational ecology have already been formalized in logic (Péli et al. 1994; Kamps and Masuch 1997; Péli and Masuch 1997; Bruggeman 1997b; Péli 1997; Hannan 1998). In terms of explanatory and predictive power, organizational ecology is one of the most advanced sociological theories. Previous logical formalizations have helped to resolve ambiguities, explicate tacit assumptions, improve the theory's logic, increase parsimony, and infer new and unforeseen conclusions. Our attempt at formalizing resource partitioning theory is part of a greater effort to formalize organizational ecology.

We chose resource partitioning theory out of socio-economical interest as well. The theory explains a peculiar phenomenon observed in a number of highly concentrated markets, *i.e.* the sudden proliferation of small specialist firms. The particular explanation provided by resource partitioning theory is novel and differs from classical economics. These qualities make a thorough investigation of the theory's explanatory argument valuable and interesting.

In our view, the key concept in resource partitioning theory, as well as in related theories, is competition. Chapter 4 of this thesis is entirely dedicated to a conceptual analysis of competition. A conceptual analysis is part of the *rational reconstruction* of a theory. Rational reconstruction, that is, determining the meaning of a theory's concepts and the structure of its argument, is the most difficult part in formalizing a theory. If a rational reconstruction is carried out well, the proper formalization, that is, the translation of the theory into a formal language, is relatively simple.

The value of a formalization attempt is not only in the construction of a sound and consistent formal theory. It is also in the usefulness of the resulting theory as an explanation for actual social phenomena. Therefore, in Chapter 5 of this thesis, we validate the formalized resource partitioning theory empirically.

With this thesis, we have the following general aims:

1. Developing heuristics for theory formalization. More specifically, developing a systematic approach to (logical) formalization of social science theory that may help social scientists to come to a better understanding of the theories in their field.

2. Applying formalization to social science theories, in particular to resource partitioning theory.

3. Clarifying social processes, in particular, competitive dynamics in the Internet search engine industry.

Overview

This thesis is organized as follows:

Chapter 2. A Logical Toolkit for Theory (Re)construction
Chapter 2 presents a systematic 5-step approach to computer supported logical formalization, which is widely applicable to sociological theory and other declarative discourse. Formalization increases rigor and precision of sociological

arguments. As a consequence, the arguments become better accessible to critical investigation, and the scientific debate is raised to a new level. The approach is demonstrated by applying it to a fragment of a theory from the sociological literature. *Chapter 2 is adapted from: Bruggeman and Vermeulen (2002).*

Chapter 3. The Logic of Organizational Markets: Thinking Through Resource Partitioning Theory

Resource partitioning theory claims that "Increasing concentration enhances the life chances of specialist organizations." By means of our 5-step formalization approach, we systematically think through this theory, specify implicit background assumptions, sharpen concepts, and rigorously check the theory's logic. As a result, we increase the theory's explanatory power, and claim— contrary to received opinion—that under certain general conditions, "resource partitioning" and the proliferation of specialists can take place independently of organizational mass and relative size effects, size localized competition, diversifying consumer tastes, increasing number of dimensions of the resource space, and changing niche widths. Our analysis furthermore shows that it is not concentration that enhances the life chances of specialists, but economies of scale. We argue that the number of organizations in the population increases—regardless of the incumbents' sizes—if scale economies come to dominate. *Chapter 3 is adapted from: Vermeulen and Bruggeman (2001).*

Chapter 4. Competition in Industries

In Chapter 4 we specify a network-representation of competition in industries. The representation consists of (1) a market network of relations between firms and environmental resources, and (2) a competition network consisting of competitive relations between firms. In a competitive relation, we distinguish the relation itself, the intensity of the relation, and the pressure that the relation imposes on the pertaining firms. By aggregating the competitive pressure exerted by its competitors, we obtain a measure for the competitive pressure on a focal firm. By aggregating the competitive pressure over all firms, we obtain a measure for industry competition. In this way, we establish a formal connection between competition on the micro level of individual firms, and competition at the macro level of industries. The micro-macro connection facilitates the connection between theories about the micro level, and those about the macro level of competition. To demonstrate the benefits of our network-representation, we measure competition at both the micro and the macro level, in the industry of Internet search engines. As an application of the model, we study the relation between competition and firm size.

Chapter 5. Distribution of Competition in The Market of Internet Search Engines

Chapter 5 investigates differentiation processes in the emerging market of Internet search engines from the perspective of organizational ecology. It uses the micro-macro link in competition, established in Chapter 4, to connect two previously unrelated fragments of organizational ecology, density dependence and resource partitioning theory. We use the combined explanatory power of both theories in an attempt to explain the emergence of Internet search engines. This emergence is characterized by processes of differentiation that are usually observed at later stages in a market. We empirically investigate competitive dynamics in the search engines market from its inception in 1993 to 2000. We find no evidence of density dependent market dynamics, but we do find evidence for resource partitioning. Furthermore, we find no evidence that market concentration and increasing market dimensionality are related to differentiation.

Chapter 2

A Logical Toolkit for Theory (Re)construction

> The very first lesson that we have a right to demand that logic shall teach us is how to make our ideas clear; and a most important one it is, depreciated only by minds who stand in need of it. To know what we think, to be masters of our own meaning, will make a solid foundation for great and weighty thought. It is most easily learned by those whose ideas are meagre and restricted; and far happier they than such as wallow helplessly in a rich mud of conceptions.
>
> —*C.S. Peirce, How to make our ideas clear.*

2.1 Introduction

[1] Social scientists communicate most of their ideas and findings in natural language. Compared to everyday conversation, though, scientific discourse is more regulated. In, for example, relating ideas to the pertaining literature, and in analyzing empirical data and displaying empirical results, authors of scientific publications commit to certain rules and procedures. As a consequence, their findings are laid open to scrutiny, criticism, and falsification by peers, who can check for themselves claims published. These self-imposed mechanisms of control and tractability distinguish scientific discourse from other kinds of discourse.

For most social scientists, their main assignment is to hypothesize about social phenomena, and to test their hypotheses empirically. Although in game theory and some other fields, hypotheses are inferred through mathematical derivations, most theoretical reasoning in the social sciences takes place in natural language.

[1]Adapted from: Bruggeman and Vermeulen (2002). Basic ideas for the heuristics presented were developed by László Pólos, Gábor Péli, Breanndán Ó Nualláin, Michael Masuch, Jaap Kamps, and the authors of this chapter, who all worked at CCSOM, now called the *Applied Logic Laboratory*, in Amsterdam.

The upside is that nearly everyone is able to understand the arguments made, or at least believes (s)he can. The downside, however, is that the flexibility of natural language comes at a cost: it is notoriously ambiguous, both conceptually and logically. Moreover, natural language has no clearcut benchmarks with respect to soundness and consistency. Consequently, a theoretical argument in natural language can easily be misinterpreted, and the logical validity of such an argument can be hard to verify, thereby challenging and sometimes violating the rules of the game. The degree of ambiguity present in social science theory would certainly not be tolerated if it would concern collecting data or analyzing empirical findings. Imagine a world void of methods and statistics, in which the researcher is left to analyze and evaluate empirical phenomena only with common sense. Nevertheless, we seem to accept such a state of affairs for our treasured theories.

In a number of recent publications, an argument has been made to use formal logic in conjunction with sophisticated computer tools, to represent and (re)construct sociological theory (Péli et al. 1994; Hannan 1998; Kamps and Pólos 1999). On the one hand, formal logic shares with mathematics rigor and precision. Moreover, it has clearcut benchmarks for soundness and consistency. On the other hand, formal logic shares with natural language, to a large extent, its sentential structure. The latter makes possible for a formalized argument to stay relatively close to its natural language counterpart, whereas mathematics often seems to represent an argument in an unrecognizable manner.

The most important reason to use formal logic is that it makes possible to reflect on scientific reasoning systematically and rigorously.[2] Logic forces the user to disambiguate the logical structure of an argument, and to lay bare each argumentative step, thereby revealing loopholes (*i.e.,* implicit assumptions), invalid inferences, and inconsistencies. The fact that in logic one can actually *prove* claims, by following a small number of clearcut rules of inference, is an advantage over informal theorizing that can hardly be overestimated.[3] On top of that, logic enables to infer new and sometimes unforeseen conclusions from established empirical facts and generalizations. If conclusions based on true assumptions are proven, they do not need empirical support in their own right, and can be transferred immediately to the set of statements we know to be true about the world.

[2]Some argue that because logical calculi are—generally—limited to two values, true and false, mathematical equations have less limitations than logical calculi have (Freese 1980, p.199). In a two-valued logic, however, one can use any mathematical function and relation, and one can reason about any mathematical equation. Moreover, just like logical statements, mathematical statements are either true or false.

[3]Although finding a proof can be an art, checking a proof object (*i.e.*, a fully written down proof) is simple, and can be fully automated. In first-order logic, which we use for our formalizations, both proof finding and proof checking can be automated, as well as model generating to check consistency.

Furthermore, logic forces to think more rigorously about the concepts that occur in an argument. Many concepts have different and colliding denotations within a field of science, or even within one theory. Also, relations between concepts may be implicitly assumed, but must be specified explicitly in order for an argument to go through.

For reducing conceptual as well as logical ambiguity, choices must be made. At least as important as a formal representation of an argument, is the explicated knowledge and motivation for the choices made along the way. Once these choices are documented, reviewers and readers who do not agree with certain propositions can trace back exactly the point where they think something might have gone wrong. Both formalization and its documented choices increase control and tractability of scientific discourse—already achieved for empirical research—and may catalyze scientific debate.

Not every argument made in the social sciences asks for rigorous logical scrutiny. Some arguments are simple and straightforward, and their logical validity is easy to establish. However, as domains described become more complex, arguments may also become more complex and harder to handle. Readers may get the feeling that there is something fishy about an argument, but not be able to put a finger exactly on the sore spot. Or, one may discuss a well-known theory with a colleague, only to discover that (s)he had a completely different understanding of it all along. In such cases, a natural tendency is to thoroughly re-investigate those parts of a publication that can be feasibly investigated, for example the statistical evidence claimed to support a theory, and to draw one's own conclusions from there.

Social science theory needs to be taken more seriously than that. It should be taken for what it is supposed to be: explanations of social phenomena, cast in logically valid arguments. Theory should be more than a context that helps to interpret correlations found in a dataset. We should start judging social science theory by its own merits, and we need a way to judge it.

This chapter attempts to provide such a way. It presents a 5-step approach to computer supported logical formalization. Our approach takes a scientific text containing an argument as a point of departure (but social theorists may take their own ideas instead), and helps to produce a formal, sound and consistent theory as a point of termination. The latter is not a termination point for theory development, though. To the contrary, a formal representation of a theory is a stepping stone for comparison, further development, and integration of theory. In this respect, the use of formal logic can contribute significantly to the accumulation and growth of knowledge in the social sciences.[4]

Our 5-step approach is designed to target formalization systematically. Each

[4] Then of course the field should not focus on problems that are mere artifacts of formalization (Hansson 2000), as happens for instance in some fields of economics.

subsequent step has the output of its preceding step as an input, and for each step, a number of heuristics (*i.e.,* tricks of the trade) is presented in this chapter, that help to gain insight in, and understanding of, a theory, its logic, and its concepts. The first three steps in our approach, constituting a so-called *rational reconstruction*, focus on reducing logical and conceptual ambiguity, by (1) identifying sentences in the text that capture the core theory, (2) analyzing and sharpening key concepts and phrases, and (3) axiomatizing informally with the aid of a conceptual model. If the rational reconstruction is done well, a (4) formalization in logic, which in turn is (5) formally tested by computer, is relatively straightforward.[5]

To illustrate the merits of our approach, we use an example from an actual social science theory. This example has the degree of ambiguity that is typical for the social sciences, and that makes a rational reconstruction difficult. The example is chosen to highlight the rational reconstruction part, since worked out examples of formal representations of sociological theories are readily available in recent literature (Péli et al. 1994; Péli 1997; Bruggeman 1997b; Kamps and Masuch 1997; Péli and Masuch 1997; Hannan 1998; Kamps and Pólos 1999; Carroll and Hannan 2000).

In sum, we present a systematic, and documenting, approach to computer supported logical formalization, involving heuristics on the one hand, and software—freely available on the web—on the other hand. After further motivating and explicating our 5-step approach in Sections 2.2 and 2.3, we apply it to a sociological example in Section 2.4. In Section 2.4, we also treat the application of specific software, *i.e.,* a theorem prover and a model generator. The heuristics are dispersed over Sections 2.2, 2.3 and 2.4. The chapter ends with a discussion and conclusions in Section 2.5.

Although this chapter can be read by social scientists with no background in formal logic, those who themselves want to formalize should acquire some knowledge of set theory (Halmos 1960; Enderton 1977) and logic (Enderton 1972; Barwise and Etchemendy 1999), in depth knowledge of the theory they want to

[5]The heuristics we use are from or inspired by mathematics (Pólya 1945), logic (Tarski 1941; Frege 1961; Quine 1986; Hodges 1998; Andréka et al. 1998), philosophy (Popper 1959; Quine 1961; Hempel 1966; Lakatos 1976), economics (Debreu 1959), game theory (Farquharson 1969), linguistics (Gamut 1991; Van Benthem 1994), computer science (Wos 1996), artificial intelligence (Kamps 1998; Kamps 1999a; Kamps 1999b), social science (Simon 1954; Coleman 1964; Blalock 1969), psychology (De Groot 1961), biology (Woodger 1937), and last but not least from the formalization projects at CCSOM (see acknowledgment). The sequence of our formalization steps is similar to approaches in computer science (Groenboom et al. 1996) and computer simulation in social science (Sastry 1997). Ideas to infer and prove theorems computationally date back centuries (Gardner 1983), although currently used theory (Beth 1962; Feigenbaum and Feldman 1963) and well developed software are more recent (Wos et al. 1991). Ideas for formalization originate in logical positivism (Ayer 1959; Neurath 1970), and the term rational reconstruction (*rationale Nachkonstruktion*) as we use it is due to Carnap (1928).

formalize, and then acquire as much experience as needed in formalization.

2.2 Rational reconstruction

In the social sciences, texts presenting theory have complex arguments stated in natural language, sometimes interspersed with graphics or mathematics. A frequently occurring problem is to find theory in these texts, and to distinguish theory from auxiliary parts, such as examples, metaphors, analogies, summaries of the work of predecessors, empirical issues, motivations, and the like. Furthermore, texts are frequently ambiguous and their arguments may have loopholes. A rational reconstruction focuses on these problems.

2.2.1 Step 1. Identifying the core theory

To dig out theory from a text, one has to know what to look for. As a benchmark, let us look at theories represented in logic. A formal theory is a set of sentences in a given formal language with an inference system; the set of sentences is "closed" under logical deduction and conclusions are validly inferred from premises according to the rules of inference.[6] This somewhat simple definition of theory, discarding intended domain, not to mention empirical and relevance criteria, has been advocated in social science by Homans (1967, 1980), among others, and suits our practical purpose fine. An important reason to choose formal logic and its definitions is that in natural language there are precise benchmarks neither for theory, nor for logical properties such as soundness and consistency.

As a heuristic to find theory in a text under investigation, and following the definition of theory, we focus on the main claims or conclusions, and subsequently on their supportive arguments. These arguments branch "upward" until no further support for the conclusion, or for intermediate conclusions, can be found in the text. The limiting case is a statement without supportive argument, *i.e.,* an argumentative "tree" consisting of only one node.

The claims and supportive arguments taken together may be considered a relevant set of sentences for a formalization attempt, and we see it as the *core theory.* The first step in our approach, then, is to identify the sentences belonging to this core theory (Fisher 1988). The remainder of the text is important too, because it indicates how the core theory should be interpreted.[7] A list of

[6] For a formal definition we refer to the technical literature (Hodges 1983; Van Dalen 1994; Van Benthem and Ter Meulen 1997) A more sophisticated view on theory, taking the dynamics of theory development into account, is in the writings of the structuralist approach (Balzer et al. 1987). See van Benthem (1982) for a broader perspective on formal theory.

[7] In some texts, one or few instances (*i.e.*, examples) of an intended theory are described in detail, but it is left to the reader to find the appropriate generalizations (Plato 1987). In some other texts, the intended theory is to be distilled from analogies or metaphors.

sentences quoted literally from the text is the output of step 1 of our approach. Along with this output, it is worthwhile to write down questions about the text on first or second reading, when looking at it with a fresh eye, and to see whether later in the formalization process they can be answered. For theory builders, it is obvious which shortcuts they can make in step 1.

Posing questions to a core theory, or to a text at large, usually points out a great deal of ambiguity. Ambiguity leads to a combinatorial explosion of readings, as our example in Section 2.4 will point out. This is fine for poetry but dangerous for scientific theories. "The implication is clear: those of us doing verbal theory in sociology need to get beyond ancestor worship and political posturing and begin the hard work of making our ideas clear enough to profit from formalization" (Kiser 1997, p.154). We distinguish two kinds of ambiguity: conceptual, to be dealt with in step 2, and logical, to be addressed in step 3.

2.2.2 Step 2. Analyzing key concepts

To prevent a plethora of readings from a set of sentences, the key concepts and phrases should be disambiguated. Analyzing and sharpening key concepts in the core theory is the second step in our approach. For each concept (or phrase) the formalizer must find out what the objects are the concept is about, what properties the objects have, and in what relations they stand. "Physical objects are postulated entities which round out and simplify our account of the flux of experience" (Quine 1961), which can also be said about sociological objects; they are not analyzed beyond the conceptualization in question. To paraphrase a well–known example: if we want to explain that Socrates is mortal, and we know that men are mortal, it helps to know the fact that Socrates is a member of the set of men. The facts that Socrates lived in Athens in the 5th century BC, and that his wife had a quick temper, can remain unexplored.

Elaborations of concepts can sometimes be found in the source text, and in other cases the reader is thrown back on other resources. Recourse may be taken to other writings of the same authors, to the authors in person, to standard textbooks, or to accepted wisdom in that particular branch of science. Furthermore, looking for relations between key concepts, which can be tacit in the source text, can yield important additional information to reconstruct an argument.

The output of step 2 is a "dictionary" of key concepts. Along with disambiguation, a dictionary should increase the parsimony of the theory by relating concepts to each other, if possible. The formalizer should try to define as many as possible concepts in terms of as few as possible "primitive" (*i.e,* undefined) concepts. In order to decide whether or not a concept can be left undefined, the following can be applied as a rule of thumb: within the set of undefined concepts, no concept should be a synonym, an element of, or a subset of another concept in the set.

A dictionary reduces conceptual ambiguity, but it does not illuminate the logical structure of the core theory. One wonders whether conclusions are sufficiently supported, or whether there are tacit background assumptions or flaws in the argument, and whether there are redundant assumptions that may be deleted.

2.2.3 Step 3. Informal axiomatization

In the third step, the line of argumentation is analyzed. The goal is to represent the core theory as a set of relatively simple sentences, with a clear logical structure. To achieve soundness, the sentences in a core theory should match each other, by allowing synonymous concepts and phrases to match. Therefore frivolous requests for stylistic variation should be temporarily suspended, and the concepts as defined in the dictionary should be implemented all through the core theory. Then, complex sentences of the core theory are broken up into simpler ones.[8] If the sentences describe certain related events (or changes), as explanatory theories do, the logical structure of each individual sentence can be clarified by taking the events (or changes) described in the sentence, and connecting them explicitly by the logical connectives ("... *and* ...", "... *or* ...", "*if* ... *then* ...", and "... *if and only if* ..."; furthermore, there is the logical negation, "it is *not* the case that...").

When logically relating the events, usually logical ambiguity shows up, and sometimes a great deal of it.[9] Contrary to the problem of conceptual ambiguity, the problem of logical ambiguity has received little attention in the social sciences, whereas even in relatively simple sentences, common sense "logical" thinking may easily fall short (Young 1988). To appreciate the difficulties posed by logical ambiguity, one should realize that for n events described in a core theory, $2^{(2^n)}$ logical sentences can be formed—if the discursive theory does not impose restrictions (see Appendix I). So if a sentence describes three events, not clearly related by the author, the formalizer has to choose the representation that best covers the intended meaning of the original sentence out of $2^{(2^3)} = 256$ possible readings.

An important category of mistakes due to logical ambiguity is confusing causal statements and conditional (*i.e.,* "if ... then ...") statements, or confusing the latter and "when ... then ..." statements. A logical consequent and a causal

[8] On the one hand, oversimplification should be avoided when discourse is disambiguated, but on the other hand "formal theories can support delicate structures that would be much more difficult to uphold and handle in the less unambiguous setting of an informal language" (Hansson 2000, p.166).

[9] According to Popper (1959), weak, *i.e.,* permissive, assumptions are to be preferred above stronger, *i.e.,* more prohibiting, versions of the same assumptions, and strong theorems are to be preferred above weak ones, in order to increase the explanatory power of the theory. Following Popper's argument, those readings should be preferred that contribute to the explanatory power of the theory.

consequence are not necessarily related, and a logical implication does not necessarily describe a sequence of events.[10] Section 2.3 will present a more systematic treatment of logical ambiguity.

Once logical ambiguities have been resolved, or the number of alternatives reduced to a feasible number, then for each resulting sentence, its role in the argument is tagged. These roles can be premise, or conclusion. A major conclusion is called a theorem, an intermediate conclusion a lemma. Premises can be assumption or definition. If a background assumption is added to fill a loophole in the argument, it is a premise too.

To keep track of the logical relations between the premises and conclusions, and to spot gaps in the argument where one expects relations, a diagram (or any other model, Spencer Brown 1957, pp.5-14) is a useful device. If loopholes in the argument show up in the diagram, the sources for filling them are the same as for the concepts in the previous step.

The modified core theory plus added background knowledge and missing assumptions is the informal axiomatization, the output of step 3, which completes the rational reconstruction. This part is far more difficult than the formalization proper, and requires the inventiveness and imagination of the formalizer to make appropriate and well argued decisions in the reconstruction.

A rational reconstruction might appear to be firm ground to evaluate soundness and consistency, in particular with the aid of a conceptual model. Formalizations of several social science theories have pointed out, however, that informal scientific arguments may exhibit logical flaws and loopholes even after a rational reconstruction.

2.3 Formalization

To overcome the shortcomings of informal theory, the set of sentences that resulted from the rational reconstruction is represented in formal logic, and the formal representation is tested for logical properties.

2.3.1 Step 4. Formalization proper

For formalization, it is best to use a logic as simple as possible. Standard, *i.e.,* first-order logic (Hodges 1983) turned out to be strong enough in all cases we are familiar with, is intuitively straightforward, and there exists useful and well tested software for it (see step 5). Although probably most scientific theories can

[10] "When ... then ..." can be regarded as a conditional statement restricted to a time point or interval. If for all time points t it holds that *if A at t, then B at t*, we can also say *when A then B*.

be formalized in first-order logic (Quine 1987 1990, p.158), sometimes other logics can be more convenient, for instance to express intentionality or belief revision (Gabbay and Guenthner 1983 1989; Van Benthem and Ter Meulen 1997). If a more sophisticated logic is chosen, the advantages of intuitive tractability and computer support go by the board, while (tacit) ontological assumptions might slip in. Moreover, the purpose of logical formalization should be to increase comprehension, not to create complexity for its own sake. If standard logic does not seem to work, it is best to consult an expert logician first. We have seen several novices blaming standard logic for their own misunderstandings of it, and for their ill-performed rational reconstructions too. For examples of theories from various disciplines represented in first-order logic, see Kyburg (1968).

2.3.2 Step 5. Formal testing

When all statements of the core theory are represented formally, attempts should be made to prove the theorem candidates. Proving by hand helps to achieve a higher level of understanding of the theory and its logical structure, and logical problems can be discovered and repaired; shortly, "improving by proving" (Lakatos 1976, p.37). But first, the formal representation should be consistent. The reason for this order is that (in standard logic) from falsehood everything follows. An inconsistent theory is therefore automatically sound, but only those theorem candidates ought to be valid that follow from a consistent set of premises.

If a theory is inconsistent, *i.e.,* if it says that both ϕ and not-ϕ are true, it can not describe any possible state of affairs in the world, and can not have a model (an instance of the theory) too (Chang and Keisler 1990).[11] This also means that an inconsistent theory can empirically be neither supported nor rejected, and empirical research is futile. Since the target theory should be consistent, one has to show that it has at least one model in which all sentences are true. As a matter of fact, to show that there is a model for the set of premises is sufficient, because then also the premises and the statements that logically follow from the premises have a model, due to the completeness of standard logic (Chang and Keisler 1990).

For first-order logic, a computer can evaluate soundness and consistency, and avoid human error. To produce a formal model of a theory, one can use an automated model generator, such as MACE,[12] and let it run on the set of premises. MACE will try to construct a model for the set of premises by assigning objects to variables and functions, and by assigning truth values to relations. The user should let MACE look for simple models first, that is, models with as few as possible objects. If MACE fails, models with more objects should be sought after, which increases computational complexity considerably, though.

[11] For simple examples of models of theories, see any introduction to logic.

[12] Both automated model generator MACE and automated theorem prover OTTER can be downloaded from `http://www-unix.mcs.anl.gov/AR/otter/`

It is important to stress that logical formalization in most cases cannot demonstrate the inconsistency of a social science text, only that some readings of an ambiguous text are inconsistent. The appropriate action after a failed consistency–check is always to reconsider one's formalization, rather than rejecting the original theory. Formalizing, then, is not establishing the ultimate reading, but trying to establish a well argued reading, thereby making choices explicit, facilitate empirical testing, and raise the level of discussion (Suppes 1968).

One can test the derivability or soundness of theorems using an automated theorem prover, such as OTTER (Wos et al. 1991). The theorem prover is given a set of (non-contradictory) premises, and the negation of the theorem to be proven. If the theorem prover finds a contradiction, then the negated theorem is false. *Ergo*: the theorem is true. Again, if this test fails, the formalizer may have to back-track to earlier steps, or may have to repair his/her own mistakes.

A theorem prover does not only tell whether or not a theorem can be proven, it actually gives a formal proof (although hard to read from the output file). Using this information, the formalizer can see which theorems build upon which premises, which in turn elucidates the argumentative structure of the theory. Moreover, it may turn out that in the derivation of the theorems, certain premises have remained unused. For parsimony, they may be omitted from the formal representation.

Finally, a theorem prover can give valuable information regarding consistency in cases where for reasons of computational complexity, MACE is not able to produce a model. This can be done by attempting to derive a nonsensical theorem, of which one is sure that it should not follow from the premises (for example $\forall x \ Nonsense(x)$, which means, for all x it holds that x is nonsense). If the proof attempt succeeds, the set of premises in all likelihood is inconsistent.

2.3.3 New results

If the theory is consistent and sound, then it makes sense to test it empirically. Although empirical testing is beyond the scope of this chapter, it is related to formalization in at least two ways. First, operationalization should be facilitated by the conceptual clarity provided by step 2. Second, the conceptual model, or another intermediate result in the formalization process, may suggest new theorems. These new theorem candidates can also be formalized and formally tested, and the current theory can be extended if the new theorems are formally true. Moreover, these extensions of the theory are additional input for empirical research, or may support formerly unexplained empirical findings.

When presenting formal results to an audience untrained in reading formulas, a summary in natural language is helpful, and an enigmatic style of presentations is always to be avoided (Hansson 2000). In our formalization papers, we also accompanied each formula with an English phrase that captures the nitty gritty.

2.4 The 5-step approach at work

To illustrate our 5-step approach, we use an example sentence. Any declarative sentence would do, but we draw from our experience in organizational ecology (Carroll and Hannan 2000).[13] In this particular theory, social organizations are seen as inert, which means that most of them cannot adapt readily and timely to their environment (Hannan and Freeman 1984). Organizational ecology studies the dynamics of populations of organizations from a Darwinian selection perspective.

2.4.1 Identifying the core theory

In resource partitioning theory, an important part of organizational ecology, we had seven sentences in the core theory, of which the first—quoted literally from the source text (Carroll and Hannan 1995, p.216)—says that:

> Early in these markets, when the arena is crowded, most firms vie for the largest possible resource base.

Questions that one may come up with when reading the sentence are: What is the state of affairs "early" in these markets? What is a "market" (a set of firms, a set of resources, or both, or perhaps something else)? What is an "arena" (perhaps a synonym of market)? What does "crowding" mean (perhaps strong competition)? Does crowding have an ordering relation (nominal, ordinal, interval, ratio)? How many is "most" with respect to the number of firms that do not "vie for the largest possible resource base"? What do the latter firms do? What makes a resource base the "largest possible" (limited competencies, the number or size of competing organizations, the limited amount of available resources, or all of these factors)?

Without a great deal of effort, to all of these questions we found several plausible answers. These answers amounted to 48 conceptually different readings of the phrase "...most firms vie for the largest possible resource base," which by no means exhaust all possible readings (see Appendix I). A more economic way to deal with the core theory is to pin down the meaning of key concepts and phrases first, rather than studying all these different readings at length.

[13]Although we use a sentence from organizational ecology as a working example, this chapter is intended neither to criticize nor to contribute to this theory. A fully worked out formalization of this theory fragment can be found in Chapter 3.

Although biologically trained formalizers might feel tempted to apply differential equations to model population dynamics in general, and resource partitioning theory in particular, we will not. Our aim is to stay as close as possible to the source text, whereas introducing mathematical models from another discipline also brings in background assumptions that are possibly not supported by the authors of our source text. New theory construction, inspired by but not based upon a source text, is a different enterprise than pursued in this case study.

2.4.2 Analyzing key concepts

In sharpening the key concepts that occur in the example sentence, we use additional information from the source text and related writings in organizational ecology. At this point we make only a basic analysis of the pertaining concepts. For a more comprehensive analysis of the concepts of resource partitioning theory, see Chapters 3 and 4.

Each population is associated with a *resource base*, a set of resources. Individual firms tap their resources from subsets of the resource base. These subsets are called *niches*.[14] The extent to which the niches of two firms overlap (*i.e.*, between 0 to 100 percent of common resources for which they compete) determines the intensity of the *competition* between these firms. In Figure 2.1, two organizations, A and B, compete for resources in the same resource base.

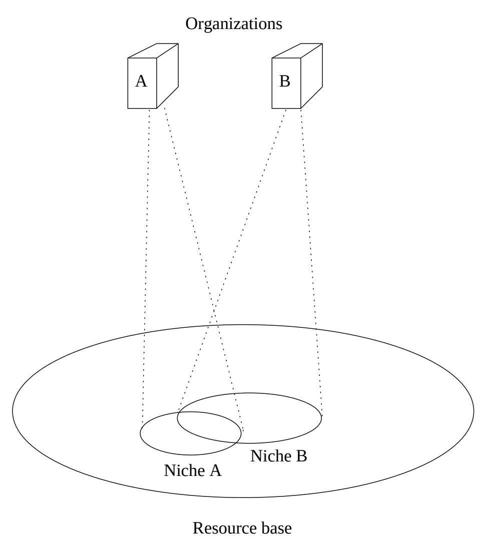


Figure 2.1: A resource base and two competing organizations

A population of firms in the domain of resource partitioning theory contains *generalist* and *specialist* firms. Organizations that appeal to a wide range of

[14]For the example here, we do not need to digress into the distinction often made between fundamental and realized niches.

resources are defined as generalist organizations. In the example sentence, "most firms vie for the largest possible resource base". From this we opt for the reading that "early in these markets" most organizations are generalist. As there is no information with regard to the ratio of generalist and specialist organizations, we only say that at the "early" time, the population contains more generalist organizations than specialist organizations.

The source text does not provide definitions of the concepts of market, arena, and crowding. In a paper that has a co-author in common with the source text, we found a mathematical definition for crowding of a set of resources by firms. In our example sentence, the arena is crowded, and we inferred that arena denotes a set of resources from which firms tap. In other words, arena and resource base are synonymous. To keep the example simple, we do not implement the mathematical definition here, but just say that "early in these markets," crowding has the value "high."

In the last phrase of our example sentence ("most firms vie for the largest possible resource base"), the concept of resource base is used as a synonym of organizational niche.

The text states that the theory of resource partitioning applies to certain kinds of markets, characterized by economies of scale and several other boundary conditions. The concept of market thus appears to denotes those parts of the universe where the authors of resource partitioning theory intend it to apply.

The temporal reference "early" is generally a relative one. For a point in time to qualify as "early in a market," it should be later than the market's beginning, and early relative to other points in time for which we know the market to exist. So, in order to define "early," we could first define some fixed time point, and use it to define the relative time point "early." On the other hand, "early" is the earliest mentioned time in the source text, as it is the point at which the process of resource partitioning starts. Before this time nothing of relevance to the process of resource partitioning happens. Because our formalization effort aims at formalizing the theory of resource partitioning, rather than market dynamics in general, we choose the beginning to be fixed, and not relative. We call it t_0, the starting point.

The dictionary for the example sentence is shown in Table 2.1. We now substitute the dictionary in the example sentence:

> At t_0, when crowding of the resource base is high, there are more generalists than specialists in the population.

2.4.3 Informal axiomatization

In the third step, the structure of argument is investigated. The logical structure of the example sentence is by no means clear. Consider the following four plausible readings of the sentence out of a much larger number (see Appendix I):

Dictionary:
early in these markets
- starting point (t_0)
arena
- resource base
crowded
- crowding is high
resource base (here)
- organizational niche
firm that vies for the largest possible resource base
- generalist
most firms vie for the largest possible resource base
- there are more generalists than specialists in the population

Table 2.1: Dictionary for the example sentence

1. **If** it is t_0 **and** crowding of the resource base is high **then** there are more generalists than specialists in the population.

2. **If** it is t_0 **and** there are more generalists than specialists in the population **then** crowding of the resource base is high.

3. **If** it is t_0 **then both** crowding of the resource base is high **and** there are more generalists than specialists in the population.

4. It is t_0 **if and only if both** crowding of the resource base is high **and** there are more generalists than specialists in the population.

In cases of logical ambiguity, the sentence in question, or the other core sentences, or the remainder of the text, may restrict the number of readings.

In addition to the source text, one may investigate the different readings systematically by using propositional logic. Although this logic is generally too simplistic to well represent scientific theories, it is precisely its simplicity that makes it a helpful tool in the rational reconstruction.

First, we draw a table that lists all possible states of affairs that might occur with respect to the events (propositions) described in the example statement. Next, we apply Popper's view that a statement can only add information to a theory if it is falsifiable by some state(s) of affairs. If no state of affairs can falsify a statement, the statement is a tautology, and should be omitted from the theory. Our example statement is falsified by states of affairs 2 to 4 (marked with an **F** in Table 2.2). On the basis of Table 2.2, there is a simple procedure (Gamut 1991, p.56) to arrive at a corresponding logical statement that is falsified in exactly these three cases. This statement corresponds to reading 3 above:

> If it is t_0 then both crowding of the resource base is high and there are more generalists than specialists in the population.

	it is t_0	crowding [...] is high	[...] more gen's than spec's [...]	
1.	true	true	true	
2.	true	true	false	**F**
3.	true	false	true	**F**
4.	true	false	false	**F**
5.	false	true	true	
6.	false	true	false	
7.	false	false	true	
8.	false	false	false	

Table 2.2: Falsifying states of affairs for the example sentence

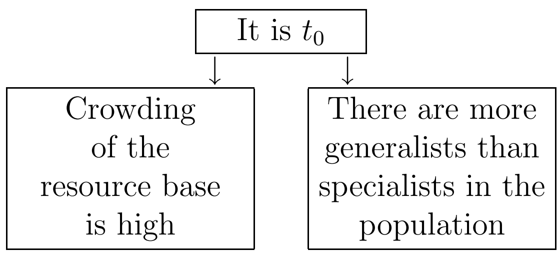


Figure 2.2: Conceptual model for the example sentence

In parallel to the informal axiomatization, a diagram or conceptual model depicting implicative logical relations between events as described by the premises often proves to be useful. Theorems and lemma's can be informally checked by tracing the implicative arrows backwards, from (desired) outcomes to their premises.

In a diagram (Figure 2.2) depicting the logical structure of our one sentence theory, the events are boxed, while arrows connecting boxes indicate implicative relations. Note again that logical relations may not coincide with causal relations or sequences of events.[15]

After analyzing the logic of the individual statements in the argument, one has to look at the logical relations between the statements, which goes beyond our one sentence example.

[15]In this specific case, the informal definition of "when ... then ..." given in footnote 10, would actually allow us to use "when ... then.." instead of "if ... then ...", because the antecedent and the consequent are both at the same time.

2.4.4 Formalization proper

The sentence that resulted from the rational reconstruction is now represented in first-order logic. First-order logic has symbols for constants, functions and relations, that the user may tailor to his/her needs. In addition to these symbols, there are variables, two quantifiers, $\forall$ (for all) and $\exists$ (there exists), and five logical connectives, $\wedge$ (and), $\vee$ (or), $\rightarrow$ (if..., then ...), $\leftrightarrow$ (if and only if) and $\neg$ (not). See Appendix II of this chapter for a more comprehensive discussion.

There are no general rules for representing informal sentences in formal logic. Only small fragments of natural language have been formalized (Van Benthem and Ter Meulen 1997). Furthermore, the representation of events described in a statement can range from one simple predicate to complicated sub-sentences. Only practice and trial and error can guide the formalizer's decisions. In our example sentence, the translation is rather straightforward.

We use a one-place relation constant SP to indicate the starting point, and a two-place relation constant RB to denote the resource base at a time. The two-place function symbol cr denotes the level of crowding of a resource base at a time. The relation symbol $High$ has one argument, and its meaning is obvious. Two one-place function symbols, n_g and n_s, denote the number of generalist firms and specialist firms in the population at a time, respectively. For the "larger than" relation, the binary relation symbol $>$ is used.

Assumption: *If it is t_0 then both crowding of the resource base is high and there are more generalists than specialists in the population.*

$$\forall t, r \qquad [SP(t) \wedge RB(t,r)] \rightarrow [High(cr(r,t)) \ \wedge \ (n_g(t) > n_s(t))]$$

Read: For all t and r, it holds that if t is the starting point and r is the resource base at t, then the crowding of r at t is high and the number of generalists at t is higher than the number of specialists at t.

2.4.5 Formal testing

Logical properties of the formal representation of the theory can now be tested. We show how consistency and soundness can be tested by computer.

Automated model generating

To check the consistency of the formal representation, we invoke MACE. For MACE input which is the same as OTTER's, see below. To see how MACE output should be read, consider the example of a possible interpretation of the $>$ relation in a model with two elements, as given in Table 2.3.

>	0	1
0	F	T
1	F	T

Table 2.3: (False) interpretation of > by MACE

The two elements of the model are named[16] 0 and 1. Table 2.3 represents a model in which $0 > 1$ and $1 > 1$ are true (T), and the other two combinations are false (F).[17] The implication is that MACE does not know how to interpret the $>$ relation symbol; although elements 0 and 1 may be ordered counter-intuitively, no ordering is able to make $1 > 1$ a true statement. To solve this problem, three meaning postulates are added that accurately define the properties of $>$:

MP 1: $\forall x \quad \neg(x > x)$

Read: For all x, it holds that it is not the case that $x > x$ (irreflexivity).

MP 2: $\forall xy \quad (x > y) \rightarrow \neg(y > x)$

Read: For all x and y, it holds that if $x > y$, then it is not the case that $y > x$ (asymmetry).

MP 3: $\forall xyz \; [(x > y) \wedge (y > z)] \rightarrow (x > z)$

Read: For all x, y, and z, it holds that if $x > y$ and $y > z$, then $x > z$ (transitivity).

Once these meaning postulates are added, the first model MACE comes up with is shown in Table 2.4.

Unfortunately, in this model there is no starting point, no resource base, crowding is not high, and there are not more generalists than specialists at any time. In this situation, wherein the antecedent of the statement is false, the statement is vacuously true. A relevant, non trivial model is wanted, not just any model. To this aim, one has to assume that there actually exists a starting point in the model, as well as a resource base. We do this by adding the following background assumption to the theory.

[16] Note that in MACE, 0 and 1 are names for elements, that might just as well have been called Abbott and Costello. In contrast to the *numbers* 0 and 1, the *elements* 0 and 1 are not related by, for example, an ordering relation.

[17] MACE tables are organized as follows: in the top left corner is either a relation or a function symbol. On the y-axis are possible first arguments of the relation or function, on the x-axis are possible second arguments. Relations between arguments are mapped to a truth value; functions of two elements are mapped to an element. The mappings are represented in the body of the MACE tables.

>	0	1
0	F	F
1	T	F

SP	0	1
	F	F

RB	0	1
0	F	F
1	F	F

High	0	1
	F	F

cr	0	1
0	0	0
1	0	0

ns	0	1
	0	1

ng	0	1
	0	1

Table 2.4: Irrelevant model of the example sentence by MACE

>	0	1
0	F	F
1	T	F

SP	0	1
	T	F

RB	0	1
0	F	T
1	F	F

High	0	1
	T	T

cr	0	1
0	0	0
1	0	0

ng	0	1
	1	0

ns	0	1
	0	0

Table 2.5: Relevant model of the example sentence by MACE

BA: $\exists tr\ SP(t) \wedge RB(t, r)$

Read: There exist an t and a r, such that t is the starting point of resource partitioning and r is the resource base at t.

With this formula added, one of the models MACE comes up with is shown in Table 2.5: In this model, if t is assigned to object 0, then t is the starting point; r is the resource base at $t = 0$ if it is assigned to object 1. Furthermore, $cr(r, t)$ is 0, and 0 is $High$; $n_g(0)$ is 1, and $n_s(0)$ is 0, hence $n_g(0) > n_s(0)$.

Proving consistency for a single or a few sentences, as we did in this example, is generally not hard. In this case, MACE provided 1536 models. If more sentences are added to the theory, the number of possible models with the same (low) cardinality usually decreases, and consistency may get harder to prove. If within a given cardinality no models are found, the formalizer switches to a higher cardinality, and has MACE try to find models there.

It is possible that the tested theory, or a particular reading of it, is consistent but that the task of model generating is too complex for the computer. To stay on the safe side, one may then attempt to demonstrate inconsistency directly, which is a simple task for a computer if the theory at hand is inconsistent indeed. For this purpose, an automated theorem prover is well suited.

Automated theorem proving

Once one or more non-trivial models of a theory are found, the formalizer may call upon OTTER to check the theory's soundness. The theorem prover is given a consistent set of premises and the negation of a theorem candidate. If the theorem prover finds the negated theorem candidate to be inconsistent with the set of premises, the theorem is sound.

Although logic is the science of reasoning, in a one-sentence example theory not much reasoning is going on.[18] To demonstrate how OTTER can be used, we give it our example statement, add the antecedent of this statement to the set of premises, and have OTTER derive the consequent. After going through this kindergarten example, the reader may try for him–/herself to have OTTER derive more exciting theorems from a different set of premises.

In OTTER, the logical conjunction ($\wedge$) is represented by `&`, the disjunction ($\vee$) is represented by `|`, the implication ($\rightarrow$) by `->` and the negation ($\neg$) by `-`. The quantifiers ($\forall$ and $\exists$) are represented by `all` and `exists`, respectively.

OTTER's input looks as follows:

```
% Meaning Postulate 1
all x ( -(x > x) ).
% Meaning Postulate 2
all x y ((x > y) -> -(y > x)).
% Meaning Postulate 3
all x y z ( ((x > y) & (y > z)) -> (x > z) ).

% Background Assumption
exists t r ( SP(t) & RB(t,r) ).

% Assumption
all t r ((SP(t) & RB(t,r)) -> (High(cr(r, t)) & (n_g(t) > n_s(t)))).

% Negation of conclusion
- (exists t, r ( High(cr(r,t)) & (n_g(t) > n_s(t)) )).
```

The first three lines of the input are the meaning postulates that describe the properties of symbol $>$. The fourth line is the background assumption stating that there actually is a t which is the starting point and that at this starting point, r is the resource base. The fifth line is the OTTER representation of the example statement, and the last line is the negation of the conclusion that crowding at t is high and there are more generalists than specialists.

[18] Formal inferencing, in contrast to rational reconstruction, is already treated in the literature extensively. For sociological examples, see for instance (Péli, Bruggeman, Masuch, and Ó Nualláin 1994; Péli 1997; Péli and Masuch 1997; Kamps and Pólos 1999).

OTTER establishes a proof in less than 0.01 seconds, and gives the following output:

```
-----> EMPTY CLAUSE at 0.00 sec ----> 12 [hyper,10,6,11] $F.

Length of proof is 2.  Level of proof is 1.

---------------- PROOF ----------------

4 [] -SP(x)| -RB(x,y)|High(cr(y,x)).
5 [] -SP(x)| -RB(x,y)|n_g(x)>n_s(x).
6 [] -High(cr(x,y))| -(n_g(y)>n_s(y)).
8 [] SP($c3).
9 [] RB($c3,$c2).
10 [hyper,9,5,8] n_g($c3)>n_s($c3).
11 [hyper,9,4,8] High(cr($c2,$c3)).
12 [hyper,10,6,11] $F.

------------ end of proof -------------
```

In steps 4 to 9 of the proof, OTTER rewrites the set of statements into the so–called "disjunctive normal form" (Fitting 1996). In next steps, OTTER applies "hyper-resolution," a logical inference rule, to the rewritten statements until in step 12 the "empty clause" is derived. This means that an inconsistency is found, so the conclusion is sound.

When using an automated theorem prover, avoid unnecessarily complex formulas, and formulas that are not necessary in a particular proof. Try to restrict the ranges of quantifiers if possible. If, for example, a property $G(x)$ is only true for organizations, $O(x)$, in the domain, then stating this fact formally, $\forall x(G(x) \rightarrow O(x))$, helps the theorem prover to shorten its proof trace. Following these guidelines can in some cases bring a proof that initially exceeds the memory capacity of the computer within a feasible range, or within the range of patience of the formalizer.

Notice again that neither an automated model generator nor a theorem prover has common sense knowledge. As we showed in the example, the human formalizer has to define the "larger than" symbol, among others.

2.5 Discussion and Conclusions

Systematic theory improvement, addressing logical and conceptual ambiguity and its underlying problems, had its adherents in the sixties (Hage 1965; Stinchcombe 1968; Blalock 1969). Since the seventies, though, such attempts have been largely abandoned, and a great deal of sociologists now look for sophistication in data collection and statistical modeling rather than theory building (Hage 1994). "Our graduate students spend years learning formal data analysis, but most do not even spend a day studying formal logic or mathematical theory" (Kiser 1997, p.153).

In our view, it makes sense for graduates to spend some time studying logic (and mathematics, for that matter). With a working knowledge of logic, and with a systematic approach to formalization and a computer at hand, logical and conceptual problems can be fruitfully addressed, and consequently be solved. A well documented formalization enhances opportunities for critical investigation of scientific arguments, and may thereby catalyze cumulative theory development. As turned out in our formalization experience, however, there are no simple tricks to translate a text presenting theory into a formal representation. Background knowledge of the theory and its intended domain—which may be tacit in the text—and careful conceptual as well as logical considerations are necessary first. Formal logic can subsequently add precision and rigor in the next step.

To achieve precision and rigor, mathematical modeling (Rapoport 1959; Coleman 1990) and computer simulation (Sastry 1997) are better known in the social sciences than logical formalization is, and they nicely complement the latter. They are well suited to model and analyze social processes, in particular complex ones with many interacting variables (Axelrod 1997; Gilbert and Troitzsch 1999).[19] Formal logic, on the other hand, is better suited to analyze complex *theories*, as logic can be seen as critical reflection on reasoning, defining, and computing (Barendregt, informal talk, 1995). Moreover, logic does not require to impose strong assumptions about metrics (Péli 1997; Hannan 1997), and better fits qualitative reasoning.

To refute a theory, or at least one assumption in it, one model as counter-example is sufficient, *i.e.*, a model in which a conclusion is false and its assumptions believed to be true. One model is also sufficient to show the consistency of a theory. For consistency and refutation, an automated model generator is therefore equally useful as a computer simulation or a mathematical model (Kamps 1998; Kamps 1999a; Kamps 1999b). It goes without saying that along with analyzing

[19]Applied to texts presenting theory, mathematical modeling and computer simulation can show that in *some* models, both assumptions and conclusions are true (provided the text provides sensible theory). If the conclusions of a theory are inferred logically, then in *all* models where the assumptions are true, the conclusions must also be true (Kamps and Pólos 1999). Notice that both mathematics and declarative text can be represented in logic. Mathematics (other than logic) is precise but not very formal, because proofs are rarely formalized to a degree that a computer can check them.

existing theories, all three strands of formal techniques can help developing new theories.

In sum, our formalization approach consists of five steps, each with a specific in- and output, and documentation of choices made along the way. It is widely applicable to the social sciences and other declarative discourse. The first three steps taken together can be seen as a hermeneutic exercise, in which a better understanding of the discursive theory is obtained; this, in other words, is a rational reconstruction, which in turn is the basis for the formalization proper. The formalization process then proceeds as hi-tech hermeneutics wherein the formal representation of the theory is used to improve the comprehension of the theory and vice versa. In the last step, the formal representation is tested for logical properties such as soundness and consistency, and possibly extended by new theorems, whereas redundant premises are deleted. New results can be obtained not only in the last step, because in each step, something about the theory can be learned.

Appendix I. Example sentence

The example sentence says that “Early in these markets, when the arena is crowded, most firms vie for the largest possible resource base.” We present different readings of this sentence, and start with the phrase “. . . most firms vie for the largest possible resource base.”

1. . . . most organizations appeal to the largest resource base possible, given the competitive forces in the population.
2. . . . most organizations realize the largest resource base possible, given the competitive forces in the population.
3. . . . most organizations appeal to the largest resource base possible, given their core competencies.
4. . . . most organizations realize the largest resource base possible, given their core competencies.
5. . . . most organizations appeal to the largest resource base possible, given the size of their fundamental niche.
6. . . . most organizations realize the largest resource base possible, given the size of their fundamental niche.
7. . . . most organizations appeal to the largest resource base possible, given the size of the resource base.
8. . . . most organizations realize the largest resource base possible, given the size of the resource base.
9. . . . most organizations attempt to appeal to the largest resource base possible, given current competitive forces in the population.

10. …most organizations attempt to realize the largest resource base possible, given current competitive forces in the population.
11. …most organizations attempt to appeal to the largest resource base possible, given expected competitive forces in the population.
12. …most organizations attempt to realize the largest resource base possible, given expected competitive forces in the population.
13. …most organizations attempt to appeal to the largest resource base possible, given their current core competencies (strong inertia).
14. …most organizations attempt to realize the largest resource base possible, given their current core competencies (strong inertia).
15. …most organizations attempt to appeal to the largest resource base possible, given their expected core competencies (weak inertia).
16. …most organizations attempt to realize the largest resource base possible, given their expected core competencies (weak inertia).
17. …most organizations attempt to appeal to the largest resource base possible, given the current size of their fundamental niche.
18. …most organizations attempt to realize the largest resource base possible, given the current size of their fundamental niche.
19. …most organizations attempt to appeal to the largest resource base possible, given the expected size of their fundamental niche.
20. …most organizations attempt to realize the largest resource base possible, given the expected size of their fundamental niche.
21. …most organizations attempt to appeal to the largest resource base possible, given the current size of the resource base.
22. …most organizations attempt to realize the largest resource base possible, given the current size of the resource base.
23. …most organizations attempt to appeal to the largest resource base possible, given the expected size of the resource base.
24. …most organizations attempt to realize the largest resource base possible, given the expected size of the resource base.

These 24 interpretations do not address the question whether the resource base (which here refers to the organizational niche) should be interpreted (1) in terms of niche width theory (Freeman and Hannan 1983) (*i.e.,* the *diversity* of resources in the niche), or (2) as the number (or value, or volume) of resources (see also Chapter 4 of this thesis). This ambiguity doubles the number of possible interpretations, generating a total of 48. At this point we stop, and ignore conceptual ambiguities in the terms early, most, markets (referring to sets of organizations, sets of resources, or unions of both types of sets?), and arena.

What needs to be addressed next is the logical ambiguity of the sentence. Does the crowding of the arena early in the market imply a certain behavior of most organizations, or does the behavior of most organizations early in the market imply that the arena is crowded? It may also be the case that the two events occur early in the market without an implicative relation, or that the occurrence of both events implies the starting point of resource partitioning. Without much effort we found 12 plausible logical readings of the sentence. We use brackets, (, and,), to avoid ambiguity.

1. **If** it is early in these markets **and** the arena is crowded **then** most organizations [...]
2. **If** it is early in these markets **and** most organizations [...] **then** the arena is crowded.
3. **If** the arena is crowded **and** most organizations [...] **then** it is early in these markets.
4. **If** it is early in these markets **then** (the arena is crowded **if and only if** most organizations [...]).
5. **If** the arena is crowded **then** (it is early in these markets **if and only if** most organizations [...]).
6. **If** most organizations [...] **then** (it is early in these markets **if and only if** crowding is high).
7. **If** it is early in these markets **then** (the arena is crowded **and** most organizations [...]).
8. **If** the arena is crowded **then** (it is early in these markets **and** most organizations [...]).
9. **If** there are more generalists than specialists **then** (it is early in these markets **and** most organizations [...]).
10. It is early in these markets **if and only if** (the arena is crowded **and** most organizations [...]).
11. The arena is crowded **if and only if** (it is early in these markets **and** most organizations [...]).
12. Most organizations [...] **if and only if** (it is early in these markets **and** the arena is crowded).

It is important to note that the logical readings mentioned above do not only differ *syntactically* (that is, by the fact that different connectives are applied), but also *semantically* (that is, all 12 readings have a distinct logical meaning). The easiest way to see this is by subjecting the sentence to a propositional logical evaluation. First, we break up the sentence into subsentences. These subsentences ("it is early in these markets," "the arena is crowded," and "most organizations vie for the largest possible resource base") we consider to be propositions. These propositions are independent in the sense that we can imagine different domains in which they can be found either to hold or not to hold, in every possible combination. Altogether, 8 different domains can be distinguished, ranging from a domain where all three propositions hold, to a domain where all three propositions do not hold. In general, for n propositions, 2^n domains can be distinguished. To form a propositional sentence from propositions, we use logical connectives. The types of connectives we use, and the order in which we use them, determine the domains by which the sentences formed are falsified. If two sentences are falsified by a different set of domains, they are semantically different. If they are falsified by the same set of domains, they are semantically equivalent. For our three propositions, we could distinguish 8 domains. With 8 domains, we can distinguish 256 sets of domains that may falsify a sentence, ranging from the

it is	**arena is**	**most**	Falsifying domains for statements 1 to 12:											
early [...]	**crowded**	**org's [...]**	1	2	3	4	5	6	7	8	9	10	11	12
true	true	true												
true	true	false	**F**			**F**	**F**		**F**	**F**		**F**	**F**	**F**
true	false	true		**F**		**F**		**F**	**F**		**F**	**F**	**F**	**F**
true	false	false							**F**			**F**		
false	true	true			**F**		**F**	**F**		**F**	**F**	**F**	**F**	**F**
false	true	false								**F**			**F**	
false	false	true									**F**			**F**
false	false	false												

Table 2.6: Falsifying states of affairs for 12 plausible readings of the example sentence

empty set (no possible domain falsifies the sentence, *ergo*, the sentence is a tautology), to the full set (all possible domains falsify the sentence, *ergo*, the sentence is a contradiction). In general, with m domains, we can distinguish 2^m sets of domains that may falsify a sentence. So, with three propositions, we can form $2^{(2^3)} = 256$ semantically different sentences, each of them falsified by a different set of domains. For our example sentence, we found 12 of the 256 semantically different interpretations plausible.

Table 2.6 shows that indeed all presented readings have a distinct logical meaning, as they are falsified by different sets of domains. In our formalization, we chose the 7th reading. As the example sentence has—at least—48 different plausible conceptual readings and—again, at least—12 plausible logical ones, it turns out that this one sentence has got $(48 \times 12 =)$ 576 different plausible reading all together. A theory consisting of, say, 7 equally ambiguous sentences would have $576^7 = 21{,}035{,}720{,}123{,}168{,}587{,}776$ plausible readings.

Appendix II. First-order logic

Like other languages, the language of first-order logic has (1) a set of *symbols*, (2) a *syntax*, which allows the user to form valid expressions, and (3) *semantics*, which give the meaning of the expressions. Unlike natural languages, first-order logic has a formal notion of consequence. This notion is realized by (4) a *model*, which tells whether a statement is true or not, and (5) a *proof system*, which enables a (true) statement to be proven.

Symbols

In first-order logic, there are seven categories of symbols:

(1) constants	such as c		
(2) relations	also called predicates, such as R		
(3) functions	such as f		
(4) variables	such as x and y		
(5) logical operators	of which we distinguish two kinds:		
	(i) *connectives*:	negation $\neg$, conjunction $\wedge$. disjunction $\vee$, implication $\rightarrow$, and equivalence $\leftrightarrow$	
	(ii) *quantifiers*:	universal, $\forall$, and existential, $\exists$	
(6) identity	$\approx$		
(7) grouping symbols	parentheses, (), and [], and commas		

Categories (1)-(3) constitute the *non-logical* symbols, the other categories are the *logical* symbols, which are the same for each first-order language.

Syntax

The syntactic rules allow us to form valid expressions from the symbols. In first-order logic, there are two types of expressions, *terms* and *formulas.*

2.5.1. Definition. Terms:

- All variables and constants are terms.
- If f is a function symbol, and $t_1, ..., t_n$ are terms, then $f(t_1, ..., t_n)$ is a term.

Terms can be compared to words; they are the building blocks of formulas. Only *well formed* formulas can have a meaning.

2.5.2. Definition. Well formed formulas, or *wff*s:

- If R is a relation symbol, and $t_1, ..., t_n$ are terms, then $R(t_1, ..., t_n)$ is a wff. This type of formula is known as an atomic formula. If R is a relation between two terms, we sometimes use the *infix* notation, $t_1 R t_2$, rather than the *prefix* notation, $R(t_1, t_2)$.
- If t_1 and t_2 are terms, then $(t_1 \approx t_2)$ is a wff.
- If $\phi_1, ..., \phi_n$ are wff's, and x is a variable, then $\neg\phi_1$, $(\phi_1 \wedge ... \wedge \phi_n)$, $(\phi_1 \vee ... \vee \phi_n)$, $(\phi_1 \rightarrow \phi_2)$, $(\phi_1 \leftrightarrow \phi_2)$, $\forall x \phi_1$, and $\exists x \phi_1$ are wff's.

A wwf in which each variable is within the scope of a quantifier, is called a *sentence.*

Semantics

In order to determine the meaning of a formula, we need to be able to interpret the logical, as well as the non–logical symbols. The logical symbols—variables, connectives, quantifiers, and the equality symbol—have a fixed meaning, which is informally given below. Let ϕ be a formula, then:

$\neg\phi_1$	means	*not* ϕ_1
$(\phi_1 \wedge ... \wedge \phi_n)$	"	ϕ_1 *and ... and* ϕ_n
$(\phi_1 \vee ... \vee \phi_n)$	"	ϕ_1 *or ... or* ϕ_n
$(\phi_1 \rightarrow \phi_2)$	"	*if* ϕ_1 *then* ϕ_2
$(\phi_1 \leftrightarrow \phi_2)$	"	ϕ_1 *if and only if* ϕ_2
$(x \approx y)$	"	x *equals* y
$\forall x\phi_1$	"	*for all* x, ϕ_1 *holds*
$\exists x\phi_1$	"	*there exists an* x, *for which* ϕ_1 *holds*

Models

To interpret the meaning of the non–logical symbols—constants, relations, and functions—we need a *model*. A model consists of a non–empty set of objects (a *universe*) and an interpretation function (an *assignment*), which maps the non–logical symbols to elements of the universe. An example of a universe is a market, which can be regarded as a set of firms, consumers, and some auxiliary objects. Relations can be defined over the objects, such as competitive relations between firms pairwise, or a supplier/consumer relation. Possible functions are a firm's size, or a consumer's budget. The assignment's function is to map, for example, the symbol $s(c)$ to "the size of firm c."

A model determines the *truth value* of a sentence. Let $\mathcal{M}$ be a model and ϕ a sentence. Then $\mathcal{M} \models \phi$ means that ϕ is true in, or *satisfied by*, $\mathcal{M}$. Let Σ be a set of sentences. $\Sigma \models \phi$ denotes that every model that satisfies Σ, also satisfies ϕ. We say that ϕ is a *logical consequence* of Σ.

Proof systems

$\Sigma \vdash \phi$ denotes that there exists a *proof* of ϕ from Σ. That means that $\phi \in \Sigma$, or ϕ is a tautology, or ϕ can be inferred from Σ by applying some *rules of inference*. A set of inference rules—a *proof system*—can be defined that is both *sound*, such that *if* $\Sigma \vdash \phi$ *then* $\Sigma \models \phi$, and *complete*, such that *if* $\Sigma \models \phi$ *then* $\Sigma \vdash \phi$.

Examples of proof systems that are both sound and complete, include *natural deduction*, intended to emulate modes of reasoning that are natural to humans, and *resolution*, which is commonly applied by computational theorem provers such as OTTER.

For readers who want to learn more about first-order logic, there are many excellent resources (introductory: Barwise and Etchemendy 1999; a linguistic approach, introductory: Gamut 1991; an overview: Hodges 1983; advanced: Van Dalen 1994; automated theorem proving: Fitting 1996).

Chapter 3

The Logic of Organizational Markets: Thinking Through Resource Partitioning Theory

3.1 Introduction

[1] As shown in Chapter 2, logical formalization of social science theories can help to resolve ambiguities, explicate tacit assumptions, improve the theory's logic, increase parsimony, and infer new and unforeseen conclusions (Péli et al. 1994; Kamps and Masuch 1997; Péli and Masuch 1997; Bruggeman 1997b; Péli 1997; Hannan 1998). Logical formalization has recently been applied, among others, to various parts of organizational ecology. As part of this greater effort, we want to analyze and formalize the explanatory structure of one of organizational ecology's sub-theories: resource partitioning theory (Carroll 1985).

Resource partitioning has become one of the most important parts of organizational ecology (Hannan and Carroll 1992; Carroll and Hannan 1995), and deals with the population dynamics of competing generalist and specialist organizations. Its main claim is that "Increasing concentration enhances the life chances of specialist organizations" (Carroll and Hannan 1995, p.p.217). Along with the proliferation of specialists, the theory claims that, as a consequence of increasing concentration, the niche overlap and competition between specialists and generalists decreases (*i.e.,* resource partitioning). The relatively new theory of resource partitioning is particularly interesting because it stands in stark contrast with older views from industrial economics. The latter see high concentration as a barrier to entry, especially for small organizations (Barney and Ouchi 1986, p.p.373, 374). Many empirical researchers have seen fit to test resource partitioning's claims, and they found corroborating evidence for resource partitioning in the car industry (Hannan et al. 1995), banking (Freeman and

[1]Adapted from: Vermeulen and Bruggeman (2001).

Lomi 1994; Lomi 1995), newspapers (Carroll 1985), the telephone industry (Barnett and Carroll 1987) beer brewery (Swaminathan and Carroll 1995), wine making (Swaminathan 1995), auditing (Boone, Bröcheler, and Carroll 2000), hi-tech industries (Mitchell 1995; Wade 1996), and the American feature film industry (Mezias and Mezias 2000). These findings are important, because higher entry rates of new firms (here, of specialists) are often associated with innovation, increased product choice, and industry renewal (Thornton 1999).

By passing resource partitioning "through the purgatory of proofs and refutations," as Lakatos (1976) phrased it, we want to get the listed advantages of logical formalization. Furthermore, we will attempt to show the redundancy of a number of assertions from organizational ecology, which a number of theorists believe to be necessary in the explanatory argument. In sum, we will try to generalize and increase the explanatory power of the theory at hand, and use logic as our tool.

In this paper, we formalize resource partitioning theory in 5 steps, with iterations if necessary. To re-iterate our 5-step approach, first, we extract the main claims and their supportive arguments from the text (Section 3.2.1). These claims and arguments taken together, we see as the *core theory.* Analyzing and sharpening key concepts in the core theory is our second step (Section 3.2.2). In the third step (Section 3.2.3), we focus on the structure of the argument. We distinguish premises (*i.e.,* assumptions, definitions, or "background" assumptions) from conclusions, use the sharpened concepts from step 2, and informally axiomatize the core theory. Loopholes in the argument (*i.e.,* "hidden" background assumptions), we fill in the course of our analysis. To visualize the structure of argument, we draw a diagram that represents the key events and their relations as described by the premises.

The first three steps are, for short, a rational reconstruction of the theory, preparing the ground for the formalization proper in step 4. In step 5, we check the two essential logical properties—soundness and consistency—computationally (Section 3.3). Checking soundness and consistency is important because conclusions that intuitively appear to follow from premises may nevertheless be false, and intuitively implausible statements may turn out to be soundly derivable. Furthermore, a computer check may reveal additional loopholes and other flaws in the argument that have been overlooked by authors, reviewers, editors, and formalizers. The results of the formalization are presented in Section 3.4, followed by a discussion in Section 3.5.

3.2 Rational reconstruction

Resource partitioning is about competing subpopulations of specialists and generalists, and asks "under what conditions will the specialist form be viable and why?" (Carroll and Hannan 1995, p.p.215). Before we go into the source text, we explicate some conceptual background from organizational ecology (Hannan and Freeman 1977; Hannan and Carroll 1992) that is important in resource partitioning theory.

Organizations tap resources from their environment, and the set of resources they tap is called their realized niche. The set of resources an organization could potentially tap in the absence of competitors, given its technology, goals, and market strategy, is its fundamental niche. Populational niches need not concern us at the moment. Two organizations compete if and only if their fundamental niches "overlap," and overlap means set intersection. Competition increases with niche overlap (Hannan and Freeman 1989).

In organizational ecology, organizations are seen as inert, which means most of them cannot adapt flexible to their environment (Hannan and Freeman 1984). If they are founded in a particular population (a collection of organizations with overlapping niches and similar form), they are likely to stay in that population for the rest of their lives. Generalists do not become specialists or vice versa. Whereas organizational structures are inert, niches may expand, shrink or drift, due to changes in the environment, organizational change (although rare), or both.

3.2.1 Core theory

Resource partitioning theory was published in 1985 by Glenn Carroll, but our main source is a more recent and slightly expanded treatment in a textbook (Carroll and Hannan 1995, p.p.215-221). This text has 92 sentences, of which seven qualify as core theory. The remainder 85 sentences rephrase (parts of) the core theory, give examples, describe concepts, pose the research question, establish connections to other chapters of the textbook, or are additional assumptions. Of these remainder assumptions, we will try to show they are redundant.

Thus our core theory is a set of quotations from Carroll and Hannan (1995, p.p.216-217), in their order of appearance in the text. Notice that in the text they appear *next* to each other.

1. Early in these markets, when the arena is crowded, most firms vie for the largest possible resource base.

2. Competition forces each to specialize to some extent to differentiate itself, although the overall strategy adopted by most firms is generalist in nature.

3. As scale economies come to dominate, only a few generalists survive and they move toward the center of the market.

4. This lessened crowding of generalists and their move to the center opens up small pockets of resources on the periphery of the market, and it is here that specialist forms usually appear and thrive.

5. In fact, the market at this point has been partitioned into generalist and specialist resources.

6. The key predictive variable in the model is the overall level of *market concentration*. [Italics in the original.]

7. When the market is not highly concentrated, specialist organizational forms will not do as well as they do when it is highly concentrated.

Questions

With respect to the core theory, one could ask whether it can explain resource partitioning and the proliferation of specialists. Moreover, there exist six additional assertions in organizational ecology, of which it is generally believed that at least some are necessary in the explanation of (part of) the outcome. First, some have argued that large organizations exert stronger competitive pressure on small organizations than the other way around. It has been shown mathematically that this assertion added to the core theory accounts for the outcome (Bruggeman 1997a). Second, large organizations consume more resources than small organizations. Some argue that if a large generalist disbands and the volume of the resource base does not decrease, resources are freed and resource pockets are opened for several small specialists to enter. In fact, this is an almost literal reading of core sentence 4, and this reading would certainly help to explain the proliferation of specialists. Third, organizational ecology has a model of size localized competition (Hannan et al. 1990), wherein organizations of similar size compete more strongly than organizations of very dissimilar sizes. This model can account for the outflow of middle sized organizations (here, small generalists) and the subsequent inflow of small specialists.

From a parsimony point of view, either assertion would require to increase the "weight" of the core theory by an assumption about the effect of organizational size, plus the conceptual ambiguity surrounding this notion. Moreover, it would suggest that resource partitioning occurs only in domains where such an additional assumption holds.

Fourth, empirical evidence suggests that at some point, consumers develop a greater variety of tastes, leading to a larger market periphery, *i.e.,* a larger resource base for specialists where they can flourish (Carroll and Hannan 1995, p.219). In this case, explaining the proliferation of specialists is rather trivial. Fifth, Péli and Nooteboom (1999) made an analogy between niche positioning and the problem of sphere packing from physics. They claim that the resource space for specialists grows if the number of dimensions of the resource space increases. And sixth, the source text (p.p.218-9) argues that the niche width of the few surviving generalists increases, although it is not clear from the text how this effect influences the outcome.

We ask, for each of these six assertions, whether it is necessary for the explanatory argument. Our aim is to establish a parsimonious set of premises on the basis of the source text, and on the basis of closely related information from organizational ecology if necessary. If we succeed, the insight gained is that additional causes, phrased in the above assertions, are not necessary for the outcome to occur. Such insight can lead to re-interpretations of previous research and a new understanding of the process of resource partitioning.

3.2.2 Key concepts

In the second step, we analyze the important concepts occurring in sentences of the core theory, and their relations, if any.

Resource base, market, and arena

A key notion is that of resource base. The resource base of a population is the set of all resources from which the organizations in the population tap. This set is also called the niche of the population (Hannan and Freeman 1989). A resource base can be partitioned into a center and a periphery. In the center, resources are relatively abundant compared to the periphery where resources are more scarce. Center and periphery are not spatial concepts, although in some cases they may take a spatial meaning, for instance for newspapers with regionally different readers.

In our view, core sentence 1 uses the notion “arena” as a synonym for resource base; in sentence 3, 4, and 5 the synonym “market” is used. In sentence 1, 6, and 7, however, market is used as an equivalent of population. For clarity, we will abstain from arena and market as synonyms for resource base, also because in most organizational theories, market denotes a collection of resources *and* organizations, as well as their mutual relations.

Generalists and specialists

Generalists and specialists are defined in terms of “niche width,” and have wide and narrow niches respectively (Freeman and Hannan 1983; Carroll and Hannan 1995, p.p.215). Each organizational population considered in resource partitioning theory contains a generalist subpopulation and a specialist subpopulation, and organizations are either generalist or specialist. In core sentence 1, generalists are said to “vie for the largest possible resource base.” The strategy of these generalists is to include as many resources as possible in their organizational niche. Having a wide niche and aiming at the center of the resource base both contribute to the generalist strategy.

In core sentence 2, organizations specialize “to differentiate themselves.” Here, the strategy of specializing is to avoid competitive pressures by reducing the crowding (see below for a definition) of the niche. Having a narrow niche and aiming at the periphery of the resource base are ways to live up to the specialist strategy. Incorporating many resources and avoiding competitive pressures are contrastive strategies. A consequence of the generalist-versus-specialist strategies is that niches of generalists are usually more crowded than niches of specialists.

The phrase "Generalists move toward the center of the resource base" is a figure of speech which means that the niches of generalist organizations increasingly include abundant resources. Notice that for resource partitioning theory, it does not matter whether resources, *e.g.,* consumer tastes, change while generalist organizations do not, or whether generalist organizations change their position with respect to the resource base.

Economies of scale

The text is clear on this concept: "An economy of scale exists when the per-unit cost of producing a product or service declines with the number of units produced." Scale economies are a main determinant for the growth of large firms that enjoy these economies.

Crowding

The notion of crowding is not defined in the source text, but its meaning is essential in the explanatory argument. There exists one definition in organizational ecology (Podolny et al. 1996) and we will use it. Crowding is defined as the degree to which organizations tap from the same resources. If many organizations aim at a relatively limited number of resources, organizational niches are crowded. If organizations differentiate themselves and aim at different resources, crowding is low. Podolny et al. (1996, p.666) focus on individual organizations in their study, and define, inspired by McPherson (1983), a notion of organizational niche crowding as the sum of its niche overlaps by other organizations.

symbols:	$A(i)$	:	crowding of the niche of organization i
	n	:	number of organizations in a population
	$a(ij)$	:	extent to which the fundamental niche of organization i is overlapped by that of organization j

3.2.1. Definition. Niche crowding (Podolny et al.)

$$A(i) = \sum_{j=1, j \neq i}^{n} a(ij)$$

Niche overlap defined as set intersection is a mathematical and clearly defined notion that abstracts away from niche dimensions. If another organization has full niche overlap with the focal organization (*i.e.,* the set intersection equals the niche of the focal organization), then $a(ij)$ is assigned the value 1. If there is no niche overlap (*i.e.,* the set intersection is empty), then $a(ij) = 0$. So, the value of the term $a(ij)$ ranges between zero and one.

Since resource partitioning is about (sub)populations, the theory must have a definition of crowding at the (sub)population level, which can be simply constructed on the basis of Podolny's definition. The crowding of a populational resource base, is the crowding of the niches of all organizations in the population.

symbols:	cr	: crowding of the populational resource base
	$\overline{A}$	: mean crowding of organizational niches

3.2.2. Definition. Populational crowding

$$cr = \sum_{i=1}^{n} A(i) = n\,\overline{A}$$

In analogy to the definition of populational crowding, the crowding measure of the generalist subpopulation is the sum of the crowding of all generalist niches, by generalists or by specialists (which may to some extent tap the same resources as generalists). The same argument applies to the specialist subpopulation. Definition 3.2.3 and 3.2.4 give measures for generalist and specialist crowding, respectively.

symbols:

cr_g	: crowding of the generalist resource base	cr_s	: crowding of the specialist resource base
n_g	: number of generalist organizations	n_s	: number of specialist organizations
$\overline{A}_g$	: mean crowding of generalist niches	$\overline{A}_s$	: mean crowding of specialist niches

3.2.3. Definition. Generalist crowding

$$cr_g = n_g \overline{A}_g$$

3.2.4. Definition. Specialist crowding

$$cr_s = n_s \overline{A}_s$$

Because in resource partitioning theory, all organizations are either generalist or specialist, the sum of generalist and specialist crowding is equal to the populational crowding, $cr = cr_g + cr_s$.

Competition

Competition is undefined in the source text, but it is defined in Hannan and Carroll's textbooks. Like crowding (Definition 3.2.1), competition increases with niche overlap and with the number of organizations; populational competition is the aggregate of competitive ties in the population (Hannan and Carroll 1992), like crowding here. In the explanatory argument of resource partitioning, we take crowding to be equivalent to competition. In Chapter 4 of this thesis, we rigorously analyze the concept of competition, and arrive at definitions of competition that are identical to the definitions of crowding above.

Concentration

Concentration denotes, informally speaking, the degree to which the resources in a market are tapped or controlled by a small number of firms of the population in that market. Usually, concentration is defined as the ratio of the aggregate size of the (3 or 4) largest firms to that of all firms in the population (Shepherd 1987). Carroll's original paper has (1) "economic concentration" (Carroll 1985, p.1262), which presumably denotes the above definition from economics, and (2) "resource concentration" (p.1275), which probably coincides with the preceding meaning, because large organizations take more resources from the resource base than small organizations do. The source text is more ambiguous, and has (3) "ownership concentration" (Carroll and Hannan 1995, p.184), along with (4) a table in which concentration is the number of organizations operating in the same market (Carroll and Hannan 1995, Table 9-1, p.185), and there is (5) concentration of specialists and of generalists independently (Carroll and Hannan 1995, p.p.192, 216). Perhaps generalist concentration could mean the same as (6) concentration of the "general interest mass market" (Carroll 1985, p.1276). We suspend deciding upon this matter and discuss concentration later, after investigating its role in the argument.

Resource partitioning

The concept of resource partitioning denotes a decrease of the extent to which generalists and specialists tap from the same resources (Carroll and Hannan 1995, p.217), *i.e.,* decreasing niche overlap of specialists and generalists.[2] In our interpretation, the opening up of small pockets of resources in the periphery (sentence 4) is a figure of speech and not meant literally. Again, it is not important whether specialists and generalists move away from each other or resources get

[2] The notion of resource partitioning suggests a partitioning of resources in a mathematical sense, as a subdivision of a set into subsets, such that each element of the set is in one of these subsets, and in no more than one. These mathematical properties hold for the resource base, which is partitioned into a center and a periphery, but not for the niches of the specialist and generalist subpopulations, which do overlap to some extent.

partitioned between specialists and generalists. The effect of both is decreasing niche overlap of generalists and specialists, which in our view is the important point.

Resource partitioning processes start "early in these markets" (sentence 1). We will refer to this point in time as the *starting point* of resource partitioning processes. These considerations complete our analysis of concepts.

3.2.3 Informal axiomatization

We will now focus on the structure of the argument. Our goal is to informally axiomatize the core theory, by representing it as a set of relatively short statements with a clear logical structure. Synonymous concepts are mapped onto one notion, using the above analyses.

For each resulting sentence, its role in the argument is tagged. These roles can be premise, or conclusion. A major conclusion is called a *theorem* (abbreviated **Thm**), an intermediate conclusion a *lemma* (**L**). In addition to these sentences, background assumptions will be added if necessary. A premise can be an assumption (**A**) a background assumption (**BA**), or a definition (**Def**).

To keep track of the logical relations between the premises, lemmas, and conclusions, and to spot gaps in the argument where we expect relations, we draw a conceptual model (Figure 3.1), to be discussed at the end of this section.

Core sentences 1 and 2

The first two core sentences provide information about the conditions that hold "early in these markets," that is, at the starting point of resource partitioning. At that time, "most firms vie for the largest possible resource base" and "the overall strategy adopted by most firms is clearly generalist in nature." We will not use this boundary condition and show that the outcome of the process can occur for any initial ratio of generalist and specialist firms.

Furthermore, "competition forces each to specialize in some extent." A straightforward reading of this phrase would be that, after the starting point, when competition is high, organizations actively limit the crowding of their niches to avoid increasing competitive pressures. But if one assumes that literally *each* organization has this strategy, a single organization that has not would falsify the assumption. An interpretation that is more in line with the Darwinian perspective of organizational ecology is that due to high competitive forces, organizations with relatively lower niche crowding are favored by selection over organizations with higher niche crowding (and for other properties similar). This assumption is about average rather than individual organizations.

A 1 After the starting point, the mean crowding of generalist niches will not increase.

A 2 After the starting point, the mean crowding of specialist niches will not increase.

Sentence 1 states that at the starting point, "the arena is crowded." The high levels of crowding, implying high levels of competition, suggest that at the starting point of resource partitioning, the environment must be near or at its carrying capacity for the population. We verified this conjecture with the author, who confirmed that the process of resource partitioning, as well as the decline of generalists and the proliferation of specialist organizations usually starts, roughly, shortly after the growth of a population has come to a halt (Carroll, personal communication).

Core sentence 3

The third core sentence mentions three events. First, "scale economies come to dominate," second "only a few generalists survive" and third "they move toward the center of the market." Since economies of scale increase mortality rates in organizational populations (p.216), we rephrase the first part of the sentence in the following manner:

A 3 If scale economies come to dominate, the number of generalist organizations decreases.

The second part of the sentence says that the niches of the remaining generalist organizations move toward the center of the resource base:

A 4 If the number of generalist organizations decreases, the niches of the remaining generalists move toward the center of the resource base.

Core sentences 4 and 5

Sentence 4 says that "This lessened crowding of generalists and their move to the center opens up small niche pockets on the periphery of the resource base," and furthermore that "it is here that specialists often appear and thrive." The cause for this lessened crowding of generalists is indicated by core sentence 3: the decreasing number of generalist organizations, due to the dominance of scale economies. From Definition 3.2.3, however, it follows that not only the number of generalist organizations, but also the mean crowding of generalist niches determines the level of generalist crowding. Assumption 1 states that after the starting point of resource partitioning, the mean crowding of generalist niches will not increase. This allows us to propose the first part of core sentence 4 as an intermediate result:

L 1 If, after the starting point, scale economies come to dominate, generalist crowding decreases.

If generalists move toward the center, their niche overlap with specialist organizations in the periphery presumably decreases. This decrease, then, will result in a lower mean crowding of specialist niches. Lower mean crowding enhances the life chances of specialist organizations and accounts for their proliferation. In other words, if we may assume that,

A 5 If generalist niches move toward the center of the resource base, the mean crowding of specialist niches decreases.

which, after all, is in line with the opening up of small pockets of resources on the periphery, we can derive the last part of core sentence 4:

L 2 If, after the starting point, generalist niches move toward the center of the resource base, and the crowding of specialists does not decrease, the number of specialists increases.

Core sentence 5 says that at this point (later than the starting point) the resource base has become partitioned into generalist and specialist resources. Since we interpreted resource partitioning as decreasing niche overlap of specialists and generalists, we rephrase core sentence 5 accordingly.

A 6 If generalist niches move toward the center of the resource base, the mean overlap of generalist and specialist niches decreases.

Let us briefly return to the starting point and ask ourselves what happens to populational crowding afterwards. Density dependence theory (Hannan and Carroll 1992) explains that at the carrying capacity (so at all times after the starting point of resource partitioning), perturbations of populational competition (hence crowding) are dampened, and competition returns to an equilibrium state. An increase of competition causes organizational outflow, which in turn reduces competition. Conversely, a decrease of competition allows for organizational inflow, due to which competition will increase. Since the same argument holds for crowding, we specify this in a background assumption.

BA 1 After the starting point, populational crowding is (approximately) stationary.

From density dependence theory, it seems to follow that after the carrying capacity has been reached, the number of organization neither declines nor resurges. When unpacking the crowding measure (Definition 3.2.2), one notices that the number of organizations and their niche overlaps can vary independently. This independence makes possible to sidestep a possible contradiction with density dependence, which does not have niche overlap as a parameter in its models. So, when density dependence has it that at the carrying capacity, populational competition stays at the same level, the number of organizations may decrease or increase if at the same time their mean niche overlap increases or decreases, respectively.

In core sentence 4, the proliferation of specialist organizations is suggested to be a consequence of both the lessened crowding of generalists and their move to the center. Lemma 2 claims, however, that a generalist move to the center is a sufficient condition for increasing numbers of specialists. We claim that also the lessened crowding of generalists is a sufficient condition for specialist proliferation. Since after the starting point, populational crowding remains stationary (BA 1), and populational crowding is the summation of generalist and specialist crowding (Definitions 3.2.2 to 3.2.4), we can derive that,

L 3 If, after the starting point, generalist crowding decreases, specialist crowding increases.

Because after the starting point, the mean crowding of specialist niches does not increase (A 2), it must be the case that the number of specialist organizations increases.

L 4 If, after the starting point, generalist crowding decreases, the number of specialists increases.

Core sentence 4 states that the "lessened crowding of generalist and their move to the center" entail specialist proliferation. We have just argued that both events *independently* can cause the proliferation of specialists. Because specialist numbers can increase without generalists moving to the center, this increase can also be expected in populations in stable environments and with highly inert organizations that do not move.

Core Sentences 6 and 7

Sentence 6 says that "the key predictive variable in the model is the overall level of *market concentration*." Finally, sentence 7 says that "when the market is not highly concentrated, specialist organizational forms will not do as well as when it is highly concentrated." These summarizing sentences address the research question and therefore are theorem candidates.

From the text (p.216) it seems that scale economies cause concentration to increase, although it is not said explicitly. But even if our reading between the lines is correct, it is certainly not said in the source text that the "predictive variable" concentration implies any of the other events described. Therefore we can not use sentence 6 in the explanation, but claim that not concentration but economies of scale cause resource partitioning and the proliferation of specialists. To substantiate our claim, we have to prove the following theorems,

Thm 1 If economies of scale come to dominate, niche overlap between generalists and specialists decreases.

Thm 2 If, after the starting point, economies of scale come to dominate, the number of specialist organizations increases.

If we succeed to prove the two theorems, then we have shown that concentration, as well as the six additional assertions (see Section 3.2.1), are not necessary in the explanation of resource partitioning processes. A diagram or conceptual model (see Figure3.1) illuminates the relations between events as described in the set of premises. The events are boxed, while the arrows indicate implicative (not necessarily causal!) relations.

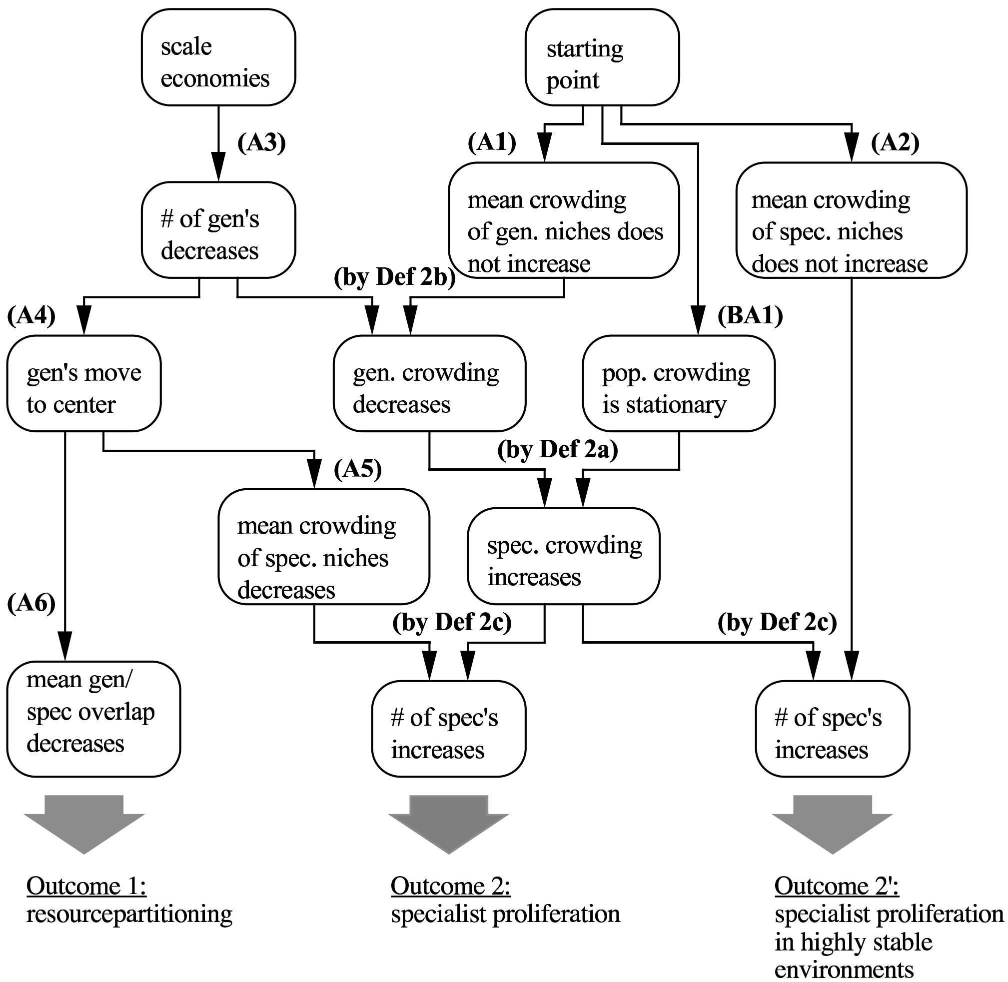


Figure 3.1: Conceptual model

We have now completed our rational reconstruction of the text, and have prepared a set of statements to be formalized.

3.3 Formalization

The set of statements is the point of departure for our logical formalization. To represent the statements formally (see Appendix), we use standard first-order logic (introductory: Gamut 1991; Barwise and Etchemendy 1999; an overview: Hodges 1983; advanced: Van Dalen 1994). First-order logic has explicit rules for constructing well-formed expressions in the language; strict rules of inference by which new expressions can be derived from existing ones; and, formal semantics by which the researcher can "see" in an exact manner what the world looks like according to a theory.

We evaluate our formal representation according to logical criteria. The formal theory should be, first, consistent and second, sound. The reason for working in this order is that according to the principle *ex falso sequitur quodlibet* (from falsehood everything follows), an inconsistent theory is automatically "sound" but not in a way any scientist would want it to be. Only those theorems should be derivable that follow from a consistent set of premises.

Consistency

If a theory is inconsistent, *i.e.,* if it says that both ϕ and not-ϕ are true, it can not describe any possible state of affairs in the world, and can not have a model in a technical sense too (Chang and Keisler 1990). Since we want resource partitioning theory to be consistent, we have to show that it has a model in which all sentences are true. To produce a formal model of the theory, we use an automated model generator, MACE,[3] which runs on the set of formal premises.

MACE produced a model of our formal representation, and we can be sure that the formal representation is consistent.

Soundness

We test the derivability of theorems and lemmas using an automated theorem prover, OTTER (Wos et al. 1991). The theorem prover is given a set of (non-contradictory) premises and the negation of the statement to be proven. If the theorem prover finds a contradiction, then the negated statement is false. *Ergo*: the statement is true.

The theorem prover confirmed that each of the lemmas and theorems is derivable from the premises, so our representation of resource partitioning theory is sound. The formally approved logical structure is depicted in Figure 3.2. For each lemma and theorem, the incoming arrows show which statements are used in its derivation. Theorem 2 also has an alternative derivation, indicated by dashed lines from Lemma 1 and 4. According to the assumptions used in this derivation,

[3] Both automated model generator MACE and automated theorem prover OTTER, which we discuss below, can be downloaded from `http://www-unix.mcs.anl.gov/AR/otter/`

specialists can also proliferate in highly stable environments with non-moving generalists (outcome 2′ in Figure 3.1).

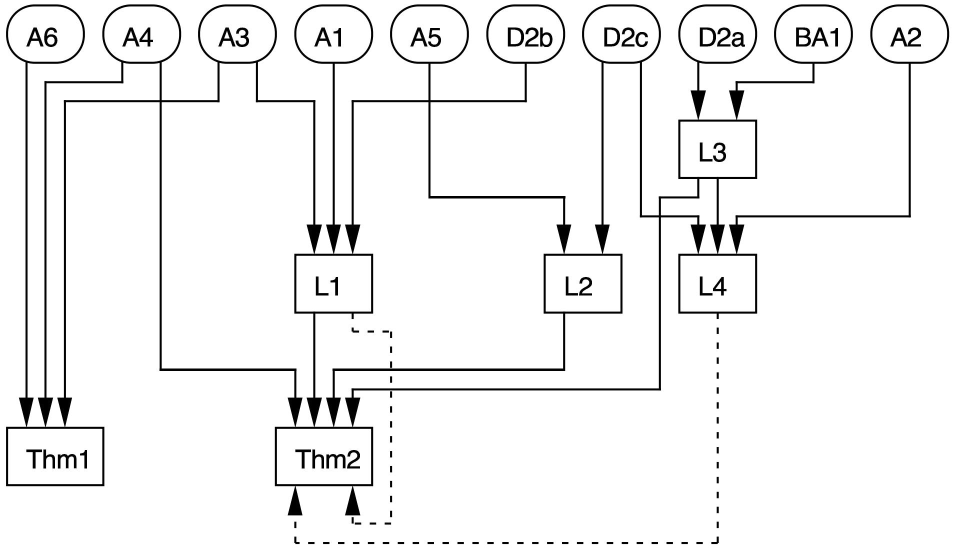


Figure 3.2: Inference structure

3.4 Results

On the basis of the logical formalization, we can now explain resource partitioning and the proliferation of specialists, in a sound, consistent, and surprisingly parsimonious way. In the explanation, we have not used any of the six additional assertions (see Section 3.2.1). They are redundant, which means that resource partitioning and the proliferation of specialists can take place without relative size effects on competition; relative size effects on the availability of resources (*i.e.,* small pockets in the resource base for specialists); size localized competition; a larger periphery due to diversifying consumer tastes; a higher number of dimensions of the resource space; and, changing niche widths of surviving generalists.

According to the source text, at the starting point, most firms pursue a generalist strategy ("vie for the largest possible resource base"). We have not used this information in any derivation, which shows that the ratio of specialists and generalists at the start does not matter for the outcome. Furthermore, we have shown that niche widths, partly characterizing specialists and generalists, are not necessary for the explanation either. Our results thus point out a far more general class of settings where the process of resource partitioning can be expected to occur.

Lemma 1–3 have demonstrated that lessened crowding of generalist niches is a sufficient condition for the increasing number of specialists, and a generalist move to the center is not necessary for the outcome to occur. The latter implies that also in stable environments and with highly inert thus non-moving generalists, specialists can proliferate. Last but not least, concentration, presented as the "key predictive variable" in the source text, plays no explanatory role. As a consequence, the two theorems answering the question "under what conditions will the specialist form be viable and why?" emphasize the role of economies of scale rather than concentration for resource partitioning (Thm 1) and the proliferation of specialists (Thm 2).

Going beyond the source text, one may ask whether these two outcomes in their turn influence the density of the population. In several empirical studies, on automobile manufacturers (Carroll and Hannan 1995, p.206) and breweries (Swaminathan and Carroll 1995, p.224) among others, the number of organizations has been found to increase unexpectedly, after a period of stability or decline. On the basis of our logical formalization, we can derive a theorem that might explain the increasing number of organizations. Due to the generalist and specialist strategies in resource partitioning theory, the mean crowding of generalist niches is higher than that of specialist niches. Furthermore, from the starting point onwards, the mean crowding of both generalist and specialist niches does not increase (A 1 and A 2). Since populational crowding is stationary (BA 1), it not only follows that the number of specialists increases (Thm 2), but also that the number of inflowing specialists is larger than the number of outflowing generalists. This means that the number of organizations in the population increases.

To end our formalization effort with this new result, we formally derive it as Theorem 3. We first assume explicitly that,

A 7 At the starting point, the mean crowding of generalist niches is higher than the mean crowding of specialist niches.

Subsequently, we derive a new lemma, starting from Definitions 3.2.2 to 3.2.4, that says that,

L 5 If the crowding of the resource base is stationary (BA 1), and the mean crowding of both generalist and specialist niches does not increase (A 1,2), and the mean crowding of generalist niches is initially higher than that of specialist niches (A 7), then the inflow of specialists is higher than the outflow of generalists.

The automated model generator confirmed that the set of premises, with Assumption 7 added, is consistent. Moreover, the theorem prover derived Theorem 3, saying that,

Thm 3 If, after the starting point scale economies come to dominate, specialist inflow is higher than generalist outflow.

This theorem is surprising because it is normally believed that for increasing density to occur, it is necessary that specialists are smaller than generalists and can thrive on less resources, or else that diversifying consumer tastes lead to a larger market periphery or to a higher dimensionality of the resource space. We have just proven that none of these assertions is necessary.

3.5 Discussion

In the long long ago, scientists took their time to study each textbook extensively. In our information-overloaded society, hardly anyone has the time for such extensive study, but still everyone expects to learn a great deal from textbooks and scientific journals. Ambiguous or sloppy discourse with many unnecessary assumptions impedes comprehension, and it is therefore important to have clear, parsimonious and logical theories with high explanatory power. With such theories, more facts can be understood on the basis of fewer information to start with. We believe that the reconstructed, and rigorously checked, theory of resource partitioning theory meets this modern demand.

A number of assertions from organizational ecology, as well as some assumptions from the core theory turned out to be redundant for explaining resource partitioning (Thm 1) and the proliferation of specialists (Thm 2), as we argued in the previous section. We have *proven* these two theorems, not just argued for them by example, metaphor, empirical generalization, or other questionable ways of reasoning that frequently pass for theory in social science (Sutton and Staw 1995). This means that the theorems and the redundancy of the listed assertions are true beyond reasonable doubt, provided that the assumptions are true, no matter how counter-intuitive the results may seem.

Logical support is far stronger than empirical support can ever be. The only empirical contribution that is on equal level is *counter* evidence against a theoretical statement (Popper 1959). In the case of resource partitioning, the published empirical results are affirmative, but confirmation does not add new explanatory information, and does not prove that the explanations are sound.

Along with logical rigor, one could ask whether our results are robust. For sure, they strongly depend on the (generalized) definition of crowding from organizational ecology. To prove the theorems, this definition, or another one wherein crowding depends on the number of organizations and on niche overlap, is necessary. The results also depend on the existence of a carrying capacity, which imposes a (flexible) upper bound on crowding. If crowding levels are too high, organizations disband or their niche overlap decreases until crowding has landed on its equilibrium level. Furthermore, if generalist niches would be *less* crowded than specialist niches (contrary to A 7), then Theorem 3 would not longer hold.

The main difference between Theorem 1 and 2 on the one hand and the source text on the other hand, is the explanatory role of economies of scale versus concentration, respectively. One could ask how important this difference is. If, on the one hand, economies of scale come to dominate, some organizations, with the highest economies of scale, grow (much) larger than the other organizations in the population, middle-sized generalists disband, and (small) specialists proliferate. Then the level of concentration increases accordingly. If, on the other hand, concentration increases, this increase might be indicative for economies of scale. If this argument is true, concentration is a proxy for economies of scale, and can be used in empirical studies. The claim that concentration is "the key predictive variable" (sentence 6) though, is not supported by the remainder of the source text. Moreover, recent empirical findings in the higher education publishing industry, are at odds with the positive correlation of concentration and economies of scale. Thornton (1999) found concentration to increase while economies of scale, and the number of foundings, decreased. Her findings seem to be inconsistent with sentences 6 and 7 from the core theory, but are consistent with our reading and formal representation.

In our formal representation, we have economies of scale at the far end of the explanation, and one may of course put a question mark right there. In some branches, such as retail, it might be the case that mergers or fusions, not treated in the current text, cause economies of scale. At the same time, anticipated economies of scale may trigger fusions and mergers, complicating the causal picture. On top of these, fusions and mergers will also increase organizational outflow directly, as well as concentration. Compared to this picture, the current formal theory is a somewhat simplified, although correct, representation of more complicated social phenomena.

According to the text, specialists attempt to have low competitive pressure, hence low niche overlap from other organizations. One could ask, what if the efforts of specialists to differentiate themselves are counteracted by generalists (or other specialists), or remain without success because of scarce peripheral resources? According to our core model (as captured by Definitions 3.2.2 to 3.2.4), the number of specialists would then not increase. This outcome is in line with industrial economics, which does not assume strategy differences between small (*i.e.*, specialist) and large (*i.e.*, generalist) organizations. Resource partitioning does assume a strategy difference, and our model shows that this difference accounts for the dissension between ecologists and economists.

One could also ask a more elementary question: what if specialists attempt to have no niche overlap at all? Without niche overlap, they would not have any competitive pressure, which would, according to organizational ecology at least, be a splendid condition for survival. Baum and Mezias (1992) and Baum and Haveman (1997) have studied the Manhattan hotel industry and argued that proximity is good for survival, because an agglomeration of hotels apparently attracts far more customers than each hotel individually would. From their studies one can infer that niche overlap is beneficial to a certain extent.

To generalize organizational ecology and to unify different strands of research, one could establish a conceptualization wherein niche overlap and the number of organizations together determine the dynamics of (sub)populations. Niche overlap acts as an "attracting" force on organizations (more resources) and at the same time as a "repelling" force (stronger competition). In such a general theory, also the effect of organizational mass should be taken into account. Barron (1999) argues convincingly that mass effects explain the decline of populational density after a maximum, a "stylized fact" that we can not explain in our current model. Our formalization of resource partitioning can be expanded with assumptions that capture such findings and considerations. The logical framework, possibly complemented by mathematical modeling or computer simulation, can provide the rigor and precision necessary for strong theory.

Thus a formal representation of a theory is not a rigid end station, but a step toward new understanding. It is superior to its natural language counterpart for thoroughly examining alternative assumptions, and provides a stepping stone for further developments of the theory. But formal theory is no panacea. Other than formal ways of thinking remain indispensable, as well as empirical research and statistical inferencing.

In the case of resource partitioning, it took us about two years to resolve conceptual ambiguities, to fill loopholes in the argument, and to make a consistent, sound, and relatively simple reconstruction of the source text. We felt like medieval hermeneuticians, equipped with high-tech automated theorem provers. Now we leave it to the reader to decide if it was worth the effort.

Appendix: Logical Representation

First-order logic has symbols for constants, functions and relations, which we introduce in Table 3.1. In addition to these symbols, logic has two quantifiers, $\forall$ (for all) and $\exists$ (there exists), and five logical connectives, $\wedge$ (and), $\vee$ (or), $\rightarrow$ (if..., then ...), $\leftrightarrow$ (if and only if) and $\neg$ (not). Fore a somewhat more comprehensive introduction to first-order logic, see Appendix II of Chapter 2.

Premisses

A 1 $\forall t_1, t_2 \;\; [\, [SP(t_1) \;\wedge\; (t_2 > t_1)] \;\rightarrow\; \neg(\overline{A}_g(t_2) > \overline{A}_g(t_1))\,]$

Read: If t_1 is the starting point, and t_2 is later than t_1, then the mean crowding of generalist niches will not be higher at t_2 than at t_1.

A 2 $\forall t_1, t_2 \;\; [\, [SP(t_1) \;\wedge (t_2 > t_1)] \;\rightarrow\; \neg(\overline{A}_s(t_2) > \overline{A}_s(t_1))\,]$

symbols (in order of appearance):

$SP(t)$	: time t is the starting point
$x > y$	: x is larger than y
$\overline{A}_g(t)$	: mean crowding of generalist niches at t (function)
$\overline{A}_s(t)$	: mean crowding of specialist niches at t (function)
$SE(t_1, t_2)$	: scale economies dominate from times t_1 to t_2
$n_g(t)$	: number of generalists at time t (function)
$GMC(t_1, t_2)$	: generalist niches move to the center of the resource base from t_1 to t_2
$\overline{a}_{gs}(t)$	: mean niche overlap between generalists and specialists at t (function)
$cr(t)$	: crowding of the resource base at t (function)
$x \simeq y$	: x is neither (significantly) larger than y, nor (significantly) smaller than y
$cr_g(t)$	: crowding of the generalist resource base at t (function)
$cr_s(t)$	: crowding of the specialist resource base at t (function)
$n_s(t)$	: number of specialists in at t (function)
$x - y$	: x minus y (function)

Table 3.1: Symbols used in the logical formalization.

Read: If t_1 is the starting point, and t_2 is later than t_1, then the mean crowding of specialist niches will not be higher at t_2 than at t_1.

A 3 $\forall t_1, t_2\ [SE(t_1, t_2) \rightarrow [(n_g(t_1) > n_g(t_2)) \wedge (t_2 > t_1)]\ ]$

Read: If scale economies dominate from t_1 to t_2, then the number of generalists at t_1 is higher than at t_2, where t_2 is later than t_1.

A 4 $\forall t_1, t_2\ [\ [(n_g(t_1) > n_g(t_2)) \wedge (t_2 > t_1)] \rightarrow GMC(t_1, t_2)\ ]$

Read: If the number of generalists is higher at t_1 than at t_2, and t_2 is later than t_1, then the generalist niches move toward the center of the resource base from t_1 to t_2.

A 5 $\forall t_1, t_2\ [GMC(t_1, t_2) \rightarrow (\overline{A}_s(t_1) > \overline{A}_s(t_2))]$

Read: If generalist niches move toward the center of the resource base from t_1 to t_2, then the mean overlap between generalist and specialist niches at t_1 is larger than at t_2.

A 6 $\forall t_1, t_2\ [GMC(t_1, t_2) \rightarrow (\overline{a}_{gs}(t_1) > \overline{a}_{gs}(t_2))]$

Read: If generalist niches move toward the center of the resource base from t_1 to t_2, then the mean niche overlap between generalists and specialists at t_1 is larger than at t_2.

A 7 $\forall t_1 \quad [SP(t_1) \;\rightarrow\; (\overline{A}_g(t_1) > \overline{A}_s(t_1))]$

Read: If t_1 is the starting point, the mean crowding of generalist niches at t_1 is higher than the mean crowding of specialist niches at t_1.

BA 1 $\forall t_1, t_2 \;\; [\,[SP(t_1) \;\wedge\; (t_2 > t_1)] \;\rightarrow\; (cr(t_1) \simeq cr(t_2))\,]$

Read: If t_1 is the starting point, and t_2 is later than t_1, then the crowding of the resource base generalist will be similar at t_1 and t_2.

We add an additional background assumption (**BA 2**), stating that a starting point of a resource partitioning process in a given population occurs only once. In other words, if a starting point has been reached, it will not be reached again. This is common sense background knowledge for humans, but not for computers.

BA 2: $\forall t_1 t_2 \;\; [\,[SP(t_1) \;\wedge\; (t_2 > t_1)] \;\rightarrow\; \neg SP(t_2)]$

Read: If t_1 is the starting point, and t_2 is later than t_1, then t_2 is not the starting point.

To be able to derive the lemmas and theorems, we also need to represent Definition 3.2.2 to 3.2.4 in first-order logic. Rather than translating the equalities straight away, we need for our derivations certain inequalities that trivially follow from this definition. For instance in the equality $cr_g = n_g \overline{A}_g$, if $\overline{A}_g$ does not increase and n_g decreases, then it is obvious that cr_g must also decrease. This inequality we call Corollary 1. In a similar way we get Corollary 2a, 2b and 3.

Cor 1 $\forall t_1, t_2 \;\; [\,[\neg(\overline{A}_g(t_2) > \overline{A}_g(t_1)) \;\wedge\; (n_g(t_1) > n_g(t_2))] \;\rightarrow\; (cr_g(t_1) > cr_g(t_2))\,]$

Read: If the mean crowding of generalist niches is not higher at t_2 than at t_1, and the number of generalists is higher at t_1 than at t_2, then the generalist resource base is more crowded at t_1 than at t_2.

Cor 2a $\forall t_1, t_2 \;\; [\,[(cr_s(t_2) > cr_s(t_1)) \;\wedge\; \neg(\overline{A}_s(t_2) > \overline{A}_s(t_1))] \;\rightarrow\; (n_s(t_2) > n_s(t_1))\,]$

Read: If the specialist resource base is more crowded at t_2 than at t_1, and the mean crowding of specialist niches is not higher at t_2 than at t_1, then the number of specialists is higher at t_2 than at t_1.

Cor 2b $\forall t_1, t_2 \ \ [\, [\neg(cr_s(t_1) > cr_s(t_2)) \ \wedge \ (\overline{A}_s(t_1) > \overline{A}_s(t_2))] \ \rightarrow \ (n_s(t_2) > n_s(t_1))\,]$

Read: If the specialist resource base is not more crowded at t_1 than at t_2, and the mean crowding of specialist niches is higher at t_1 than at t_2, then the number of specialists is higher at t_2 than at t_1.

Cor 3 $\forall t_1, t_2 \ \ [\, [(cr(t_1) \simeq cr(t_2)) \ \wedge \ (cr_g(t_1) > cr_g(t_2))] \ \rightarrow \ (cr_s(t_2) > cr_s(t_1))\,]$

Read: If the crowding of the resource base is similar at t_1 and t_2, and the generalist resource base is more crowded at t_1 than at t_2, then the specialist resource base is more crowded at t_2 than at t_1.

Lemmas

L 1 $\forall t_1, t_2 \ \ [\, [SP(t_1) \ \wedge \ SE(t_1, t_2)] \ \rightarrow \ (cr_g(t_1) > cr_g(t_2))\,]$

Read: If t_1 is the starting point, and scale economies dominate from t_1 to t_2, then the generalist resource base is more crowded at t_1 than at t_2.

L 2 $\forall t_1, t_2 \ \ [\, [SP(t_1) \ \wedge \ GMC(t_1, t_2) \ \wedge \ \neg(cr_s(t_1) > cr_s(t_2))] \ \rightarrow$
$(n_s(t_2) > n_s(t_1))\,]$

Read: If t_1 is the starting point, generalist niches move toward the center of the resource base from t_1 to t_2, and the crowding of the specialist resource base is not higher at t_1 than at t_2, then the number of specialists is higher at t_2 than at t_1.

L 3 $\forall t_1, t_2 \ \ [\, [SP(t_1) \ \wedge \ (cr_g(t_1) > cr_g(t_2)) \ \wedge \ (t_2 > t_1)] \ \rightarrow \ (cr_s(t_2) > cr_s(t_1))\,]$

Read: If t_1 is the starting point, the generalist resource base is more crowded at t_1 than at t_2, and t_2 is later than t_1, then the specialist resource base is more crowded at t_2 than at t_1.

L 4 $\forall t_1, t_2 \ \ [\, [SP(t_1) \ \wedge \ (cr_g(t_1) > cr_g(t_2)) \ \wedge \ (t_2 > t_1)] \ \rightarrow \ (n_s(t_2) > n_s(t_1))\,]$

Read: If t_1 is the starting point, the generalist resource base is more crowded at t_1 than at t_2, and t_2 is later than t_1, then the number of specialists is higher at t_2 than at t_1.

L 5 $\forall t_1, t_2 \ \ [\, [(cr(t_1) \simeq cr(t_2)) \ \wedge \ \neg(\overline{A}_g(t_2) > \overline{A}_g(t_1)) \ \wedge \ \neg(\overline{A}_s(t_2) > \overline{A}_s(t_1))$
$\wedge \ (\overline{A}_g(t_1) > \overline{A}_s(t_1))] \ \rightarrow \ ((n_s(t_2) - n_s(t_1)) > (n_g(t_1) - n_g(t_2))\,]$

Read: If the crowding of the resource base is similar at t_1 and t_2, the mean crowding of generalist niches is not higher at t_2 than at t_1, the mean crowding of specialist niches is not higher at t_2 than at t_1, and the mean crowding of generalist niches at t_1 is higher than the mean crowding of specialist niches at t_1, then the number of specialists at t_2 minus the number of specialists at t_1 is higher than the number of generalists at t_1 minus the number of generalists at t_2.

Theorems

Thm 1 $\forall t_1, t_2 \ \ [SE(t_1, t_2) \ \rightarrow \ (\overline{a}_{gs}(t_1) > \overline{a}_{gs}(t_2))]$

Read: If scale economies dominate from t_1 to t_2, then the mean niche overlap between generalists and specialists is larger at t_1 than at t_2.

Thm 2 $\forall t_1, t_2 \ \ [\,[SP(t_1) \ \wedge \ SE(t_1, t_2)] \ \rightarrow \ (n_s(t_2) > n_s(t_1))\,]$

Read: If t_1 is the starting point, and scale economies dominate from t_1 to t_2, then the number of specialists is larger at t_1 than at t_2.

Thm 3 $\forall t_1, t_2 \ \ [\,[SP(t_1) \ \wedge \ SE(t_1, t_2)] \ \rightarrow$
$[(n_g(t_1) > n_g(t_2)) \ \wedge \ ((n_s(t_2) - n_s(t_1)) > (n_g(t_1) - n_g(t_2)))]\,]$

Read: If t_1 is the starting point, and scale economies dominate from t_1 to t_2, then the number of generalists is higher at t_1 than at t_2, and the number of specialists at t_2 minus the number of specialists at t_1 is higher than the number of generalists at t_1 minus the number of generalists at t_2.

Chapter 4

Competition in Industries: A Network-Representation

4.1 Introduction

[1] In 1977, in their seminal paper "The Population Ecology of Organizations" Hannan and Freeman applied biological ecological models to organizational markets, a scientific domain previously monopolized by economists. Organizational ecology and economics had in common the use of the concept of competition, although the meaning of competition in both theoretical contexts is different.

In economics, competition is usually analyzed as a relation between individual actors. Particularly the competitive relation in a pair of firms, and the implications of the relation for both firms, have been subject to extensive analysis since Hotelling (1929). Studies that discuss competition in groups of firms, usually employ micro level analyses as well (Eaton and Lipsey 1975).

In contrast, competition in ecological models of the market has been studied almost exclusively at the macro level, that is, at the level of industries (Carroll and Hannan 2000, pp.7-8).[2] In organizational ecology, competition is regarded a property of an industry, rather than of interacting actors. One of the crucial points of organizational ecology is that organizational adaption occurs at the industry level, not at the level of firms. In other words, adaption is governed by organizational founding and mortality, not by organizational change (Hannan and Freeman 1984). Founding and mortality are assumed to be dependent on industry competition, which consequently acquired the ecologists' attention.

In a number of organizational ecological studies though, the focus has been on the relation between founding and mortality rates and (1) organizational size (Baum 1995) (Barnett 1997), (2) level of generalism (or niche width) (Freeman

[1] Empirical data used in this chapter was courtesy of Northern Light Search.

[2] Within organizational ecology, the paper by Podolny, Stuart, and Hannan (1996) is a notable exception. See Sections 4.2.6 and 4.2.7.

and Hannan 1983), (3) central or peripheral market positioning (Carroll and Hannan 2000, p.268), (4) age (Hannan 1998), or (5) similarity of resource requirements (Baum and Mezias 1992). In each of these studies, it is assumed that, within an industry, one (type of) firm can have a competitive advantage over another (type of) firm. Hence, these studies—often implicitly—assume the competitive conditions for individual firms to differ, and thus, do not assume one competitive condition to hold for the entire industry.

Some organizational ecological studies suggest the existence of a relation between micro-level competition and macro-level competition, although without specifying the relation. For example, Carroll and Hannan (2000, pp.264-267) generalize the competitive advantage of larger firms over smaller firms in pairwise competitive relations, to the group level: small size is theorized to be bad for survival. Carroll and Hannan (1989a) suggest that a firm has a competitive disadvantage, if, at the time of founding of the firm, industry competition is strong.

Understanding and appreciating these suggestions would certainly be facilitated by an understanding of how micro-level and macro-level competition are related. However, in organizational ecology, the nature of the relation between both levels of competition has never been specified.

In this chapter we aim to specify a formal micro-macro link in competition. In doing so, we follow in the tradition of Coleman (1990). We analyze the competitive interaction between individual firms first, and on that analysis build measures for competition at the industry level, which are aimed to be consistent with measures applied by organizational ecology. By linking different levels of competition, we have two objectives. First, the link should clarify the repercussions of individual competitive interactions on group level competition, and vice versa. Second, the link should facilitate the connection of ideas, theories, and models about competition at the micro level, to those about competition at the macro level.

To establish a formal link between the micro and the macro level of competition, we develop a network-representation of competitive relations between firms. There are two reasons for applying network analysis to competition. First, the micro level of competition consists of relations between individual firms; network analysis provides a framework to study (collections of) relations. Second, the aggregation techniques used in network analysis facilitate establishing the desired micro-macro link.

This chapter is organized as follows: in Section 4.2 we will start out with a simple representation that captures the basic notion of firm-on-firm competition. Subsequently, we provide a basic representation of a market, and upon that representation define a measure of intensity of competitive relations. We show that our measure of competitive intensity formalizes the crucial ecological notion of niche intersection. Using the measure of competitive intensity, we define a measure of competitive pressure. We show that by merely aggregating the competitive pressure over all firms in an industry, a measure of industry competition is obtained, which is consistent with industry competition in organizational ecology. At that point, a formal link between micro and macro level competition is established. We proceed by further exploring the basic market representation, and investigate how changing the representation affects the measure of industry competition. Subsequently, we focus on the relation between size distributions of firms and the distribution of competition over an industry.

In Section 4.3, we show that the presence of competitive relations, competitive intensity, and competitive pressure can be measured directly at the level of individual firms. As a result of the formally established micro-macro link, measurements of competition at macro level come for free. The fact that the intensity of competitive relations can be operationally defined and directly measured—also dynamically—in a relatively simple way, speaks in favor of the network-representation; the interaction between (changing) competitive relations in individual pairs of firms and (changing) industry dynamics has rarely been empirically investigated (Borenstein and Netz 1999). We empirically investigate the relation between the micro and the macro level of competition by measuring competitive relations between Internet search engines, over a period of seven years. The empirical analysis confirms our conjectures about the relation between competition, industry size, firm size and differentiation.

4.2 Industries as networks of competitive relations

The aim is to develop a network-representation of competition that (1) connects competitive relations between individual actors to industry competition, (2) quantifies both competitive relations between individual actors and industry competition, and (3) can be applied in empirical research of competitive dynamics in industries.

To reach this aim we represent competition as a network of competitive relations between firms. In Section 4.2.1 we will discuss the network itself, in the subsequent sections we will focus on the relations that constitute the network.

4.2.1 Competitive relations between firms: potential and actual

We start out with a simple common-sense approach to competition. Let us assume industry I to be a set of N firms. Assume furthermore that a pair of firms $i, j \in I$ can entertain a competitive relation, which we tentatively characterize as competition for environmental resources. Because of this competitive relation, both firms have to invest money and energy in price competition, advertising, product improvement, marketing, innovation, account managing, networking, lobbying, sponsoring, packaging, deliveries, gifts, and many other activities that may persuade consumers to choose one firm's product instead of that of its competitor.[3] The competitive relations, s_{ij}, constitute a competition graph S. In Section 4.2.3 we will characterize s_{ij}, and define S. For now, we assume s_{ij} to be symmetric, and $s_{ii} = 0$.

Because industry I has N members, each member has $N-1$ potential competitive relations, aggregating to a total of $\frac{N^2-N}{2}$ potential relations in the industry. The implication is that (1) the potential number of competitive relations in an industry increases with the number of firms at an increasing rate, and (2) the potential number of competitive relations added by new entrants in populated industries is higher than in less populated industries.

The potential number of competitive relations may differ from the actual number of competitive relations. Let $s_{ij} = 1$ if firm i and firm j compete, and $s_{ij} = 0$ otherwise. Let CR denote the number of actual competitive relations in the industry. That is,[4]

$$CR = \frac{\sum_{ij,i\neq j} s_{ij}}{2} \tag{4.1}$$

The ratio of the number of actual competitive relations and the number of potential competitive relations is defined as the density[5] of the S-graph, ΔS (Wasserman and Faust 1994, p.101). Measure CR is a function of the number of firms in the industry, and the density of the competition graph; it can also be written as:

$$CR = \frac{N^2 - N}{2}\Delta S \tag{4.2}$$

Organizational ecology assumes the presence of an actual competitive relation between all members of an industry (Carroll and Hannan 2000, p.65). By doing so, organizational ecology assumes $\Delta S = 1$, and $CR = \frac{N^2-N}{2}$. The theory's measure for industry competition is based on the assumption of full density of the competition graph; hence, the measure does not accommodate for industries in which some pairs of firms do not compete.

[3] Although in this context we refer to "industries," "firms," "products" and "consumers," the network-representation of competition that is presented here applies to all populations of organizations that provide resources in exchange for—or supported by—resources controlled by other resource providers. So, the model can also be applied to, for example, research institutes competing for funding of science funds.

[4] $\sum_{ij,i\neq j}$ is shorthand for $\sum_{i=1}^{N}\sum_{j=1,j\neq i}^{N}$.

[5] Note, that in organizational ecology, the term density is used to denote the number of firms in an industry; in this case N. We reserve the term density for the density of a graph, and will not refer to N as density.

4.2.2 Appeal to resources: a bimodal graph

We adopt the framework of organizational ecology and assume that firms tap resources from a resource base, R. Resources constitute all things that firms need in order to survive (Carroll and Hannan 2000, p.199). The most common example of a resource is a consumer, but other examples of resources are employees, loans, housing, licenses, and more (Sohn 2001). We assume resource base R to be a finite set of resources.

In order to obtain resources, firms make *appeals* to resources. We define an appeal relation p_{ir} to be a relation between a firm $i \in I$ and a resource $r \in R$.[6]

4.2.1. Definition. Appeal relation p
p is a relation between $i \in I$ and $r \in R$. So, $p \in P = I \times R$, such that $p = (i, r) = 1$ iff[7] firm i appeals to resource r. Otherwise, $p = (i, r) = 0$. We denote a tuple $(i, r) \in P$ as p_{ir}.

We define the set of resources that a firm i appeals to, to be the *niche*[8] of i.

4.2.2. Definition. Niche R_i
$R_i \subseteq R$ is the niche of $i \in I$ iff it holds that $r \in R_i$ iff $p_{ir} = 1$.

In terms of network analysis, the niche of firm i is the resource degree of i.

We assume resources to be social actors, or to be controlled by social actors. In the latter case, we assume each resource to be controlled by one social actor. Furthermore, no member of focal industry I is a resource, or controls a resource. We assume R and P to be industry specific: for all $r \in R$, there is a firm $i \in I$ such that $p_{ir} = 1$. Furthermore, it holds for all $i \in I$, that if $p_{ir} = 1$, then $r \in R$.

In our framework, the bimodal[9] graph P represents a market, as shown in Figure 4.1.

4.2.3 Competition between firms: a unimodal graph

We define a competitive relation s_{ij} between firms i and j, as follows:

[6] In the definitions below we use some set theoric notation. $x \in X$ denotes that x is an element of set X, and $X \subseteq Y$ denotes that X is a subset of set Y. We write $X \cap Y$ to denote the conjunction—or intersection—of sets X and Y, and $X \cup Y$ to denote their disjunction—or union. We use $|X|$ to denote the size—or "cardinality"—of set X, *i.e.*, the number of elements in set X. Finally, $\emptyset$ denotes the empty set.

[7] The expression "iff" denotes: "if and only if."

[8] Traditionally, in organizational ecology, the niche is defined as a attribute of an industry, rather than of a firm. McPherson (1983) convincingly argued in favor of the organizational niche. Moreover, in organizational ecology the niche is considered a much more complicated structure than just a set of resources. For a more extensive discussion, see the discussion section of this chapter.

[9] Bimodal graphs consist of relations between elements of two disjunct sets.

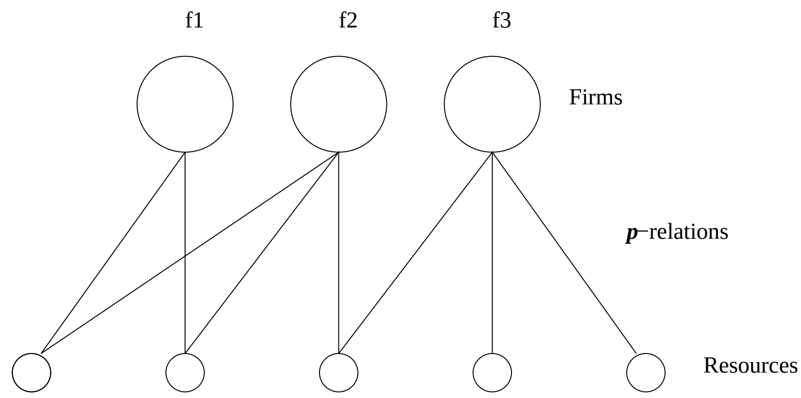


Figure 4.1: Basic market representation

4.2.3. Definition. Competitive relation s
s is a relation between $i, j \in I$. So, $s \in S = I \times I$. For all $i \neq j$ it holds that $s_{ij} = 1$ iff there exists an $r \in R$ such that $p_{ir} = 1$ and $p_{jr} = 1$. Otherwise $s_{ij} = 0$.

Simply put, Definition 4.2.3 states that two firms entertain a competitive relation if they appeal to the same resource.

In the bimodal graph in Figure 4.1, the appeals of f1 and f2 coincide, as do the appeals of f2 and f3; the appeals of f1 and f3 do not coincide. By use of Definition 4.2.3, the bimodal P-graph of Figure 4.1 translates into the unimodal[10] S-graph of Figure 4.2.

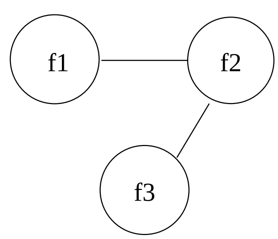


Figure 4.2: Competition graph

The current framework allows a competitive relation to either be present or not, and does not accommodate for the intuitive notion that some pairs of firms compete more than others.

[10]Unimodal graphs consist of relations between elements of one set.

4.2.4 Competition about resources: competitive intensity

Definition 4.2.3 states that two firms compete if their appeals coincide on at least one resource. Our intuition is that competition in a pair of firms can be less or more intensive. To make this intuitive notion more precise we define the notion of competitive intensity. We claim that the intensity, v_{ij}, of a competitive relation s_{ij}, in a pair of firms $i, j \in I$, is determined by the number of resources where the appeals of i and j coincide. So:

4.2.4. DEFINITION. Competitive intensity v
Let $R_{ij} = R_i \cap R_j$. Now, v is a relation between $i, j \in I$. So, $v \in V = I \times I$. For all $i \neq j$, it holds that $v_{ij} = |R_{ij}|$.

Remember (footnote 6) that $|R_{ij}|$ denotes the size—or cardinality—of set R_{ij}. So, v_{ij} is equal to the number of resources for which both $p_{ir} = 1$ and $p_{jr} = 1$. In short, v_{ij} denotes the number of resources where the appeals of firms i and j coincide. Notice that $0 \leq v_{ij} \leq R$, and $v_{ij} = v_{ji}$.

To re-iterate the example of Figure 4.1, we see that the appeals of f1 to f2 coincide on two resources, the appeals of f2 to f3 coincide on one resource and the appeals of f1 to f3 do not coincide. By use of Definition 4.2.4, the bimodal P-graph of Figure 4.1 translates into the unimodal multiple V-graph of Figure 4.3.

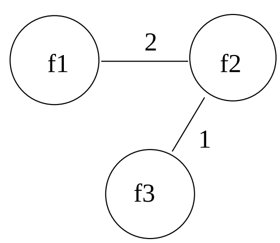


Figure 4.3: Multiple graph of competitive intensities

4.2.5 Competitive intensity versus niche intersection

We defined the niche of firm i as the set of resources that i appeals to. Figure 4.4 shows the niches in the market of Figure 4.1 as the dotted areas at the bottom of the figure. We see that the niches of f1 and f2 have an intersection of size 2, f2 and f3 have an intersection of size 1, and f1 and f3 have no niche intersection. As it turns out, competitive intensity and niche intersection are identical. More formally, $v_{ij} = |R_{ij}| = |R_i \cap R_j|$.

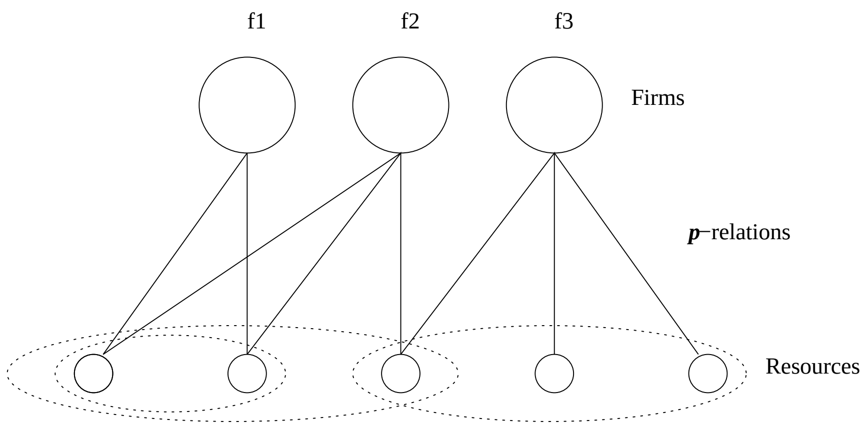


Figure 4.4: Competitive intensity versus niche intersection

4.2.6 Competitive pressure

We have now defined the presence of a competitive relation, and the intensity of a competitive relation. Note that these notions differ somewhat from organizational ecology's concept of "competition." In organizational ecology, competition is the "negative effect of the presence of one or more actors on the life chances or growth rates of some focal actor"(Carroll and Hannan 2000, p.225). In our framework, we call this negative effect "competitive pressure."

At first sight, competitive pressure may seem similar to competitive intensity: the more intense a competitive relation, the more competitive pressure the relation generates. Figure 4.5 gives an example where competitive intensity and competitive pressure indeed amount to the same. It shows two pairs of firms; one pair that makes many coinciding appeals and intuitively experiences strong competitive pressure, and another pair that makes few coinciding appeals and intuitively experiences weak competitive pressure.

However, if we consider a competitive relation between firms with different niche sizes, we see a difference between competitive intensity and competitive pressure. Consider Figure 4.6. Competitive intensity, as we know, is determined by niche intersection, which is equally large for both firms. However, only a few of firm j's appeals coincide with the appeals of i, whereas most of i's appeals coincide with the appeals of j. As a consequence, j is more of a competitive threat to i than vice versa, and intuitively the competitive pressure by j on i is larger than the other way around.[11]

To make this intuitive notion more precise, we define competitive pressure as the proportion of a firm's resources that is targeted by a competing firm. In other words, the competitive pressure that a firm j imposes on a firm i is the

[11]For an application of the notion of asymmetric competitive pressure to Burt's structural holes theory, see (Bruggeman, Carnabuci, and Vermeulen 2002).

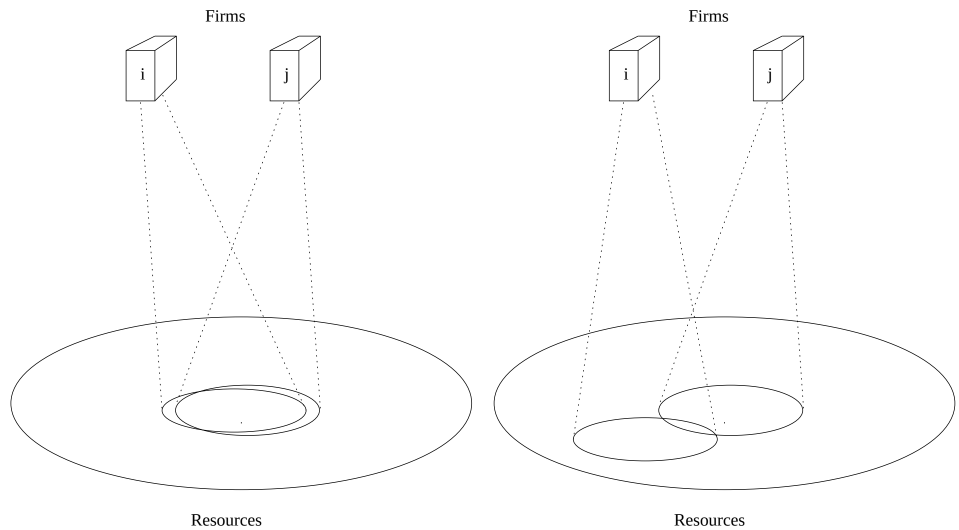


Figure 4.5: Strong vs. weak competitive pressure

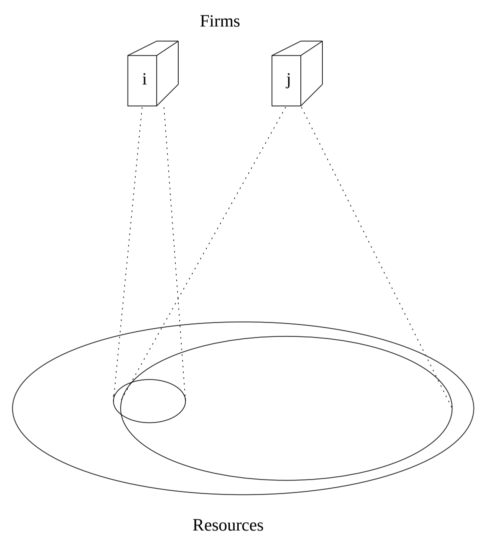


Figure 4.6: Asymmetric competitive pressure

proportion of the niche of i that the niche of j overlaps.

First, we define the size,[12] v_i, of the niche of firm i:

4.2.5. DEFINITION. Niche size v_i
v_i is the size of i's niche $R_i \subseteq R$, so $v_i = |R_i|$.

Now we define competitive pressure of firm j on firm i as follows:

4.2.6. DEFINITION. Competitive pressure j on i, c_{ij}

$$c_{ij} = \frac{v_{ij}}{v_i}$$

Note that $0 \leq c_{ij} \leq 1$. Note further that c_{ij} is not necessarily equal to c_{ji}. In terms of network theory, we say that c is not symmetric. In fact, if $v_i > v_j$, then $c_{ij} < c_{ji}$; the firm with the larger niche experiences less competitive pressure. Figure 4.7 shows the asymmetry in the competitive pressure on the firms in our previous example.

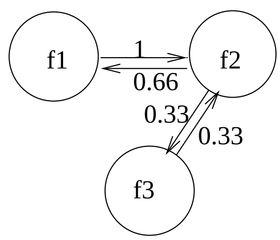


Figure 4.7: Valued graph of competitive pressures

In Section 4.2.5 we compared our measure of competitive intensity with the organizational ecological notion of "niche intersection." The measure of competitive pressure that we define in the current section is similar to Podolny et al.'s measure for "(asymmetric) niche overlap" (1996). We discussed this measure at some length in Chapter 3 of this thesis.

We have now introduced four networks representing different features of the same competitive relations: (1) the bimodal P-graph of appeal relations (Figure 4.1), (2) the unimodal symmetric S-graph of dichotomous competitive relations, (3) the symmetric unimodal multiple V-graph of competitive intensities (Figure 4.3), and (4) the directed unimodal valued C-graph of competitive pressures (Figure 4.7).

[12]Note that niche size, v_i, should not be confused with the organizational ecological term niche width, as defined in, for example, (Freeman and Hannan 1983). Niche width denotes degree of generalism, that is, the *diversity* of the resources that a firm appeals to. In our framework, there is no measure for the diversity of resources, hence, niche width is not defined.

4.2.7 Aggregating competitive pressure

We call the total competitive pressure that its competitors impose[13] on firm i, c_i.

$$c_i = \sum_{j \neq i} c_{ij} = \frac{\sum_{j \neq i} v_{ij}}{v_i} \tag{4.3}$$

Note that $0 \leq c_i \leq (N-1)$.

The measure of total competitive pressure on i is similar to Podolny et al.'s measure for "niche crowding" (1996). In Chapter 3, we assumed crowding and competition to have the same meaning. At this point, we should specify that notion somewhat: crowding means competitive *pressure*.

By adding the competitive pressure imposed on all firms by their competitors, we calculate the total competitive pressure in the industry. Let C be the degree of industry competition. Then,

$$C = \sum_{i=1}^{N} c_i$$

As,[14]

$$C = \sum_{i=1}^{N} c_i = N\, \overline{c_i} = N(N-1)\overline{c_{ij}}$$

we can also write,

$$C = (N^2 - N)\, \overline{c_{ij}}$$

Notice that $0 \leq \overline{c_{ij}} \leq 1$. As a consequence, $0 \leq C \leq (N^2 - N)$. Notice that for $N = 0$, we have $C = 0$. Moreover, for $N = 1$, we have $C = 0$, as well. The definition of C is in line with the intuition that in empty industries, as well as in industries with only one member, there is no competitive pressure.

At this point the micro-macro link between competition in pairs of firms and industrial competition is established.

4.2.8 Back to the market network: consumer preferences

In Sections 4.2.2 to 4.2.7 we focused on the competition network. Now we return, for a moment, to the underlying market network. Recall that in our framework, the market network P consists of appeal relations between firms and resources. We assumed these p-relations to be dichotomous: a firm either appeals to a resource or it does not.

We also assumed the resources to be, or to be controlled by, social actors. By representing appeals as dichotomous relations, we assume that social actors

[13] $\sum_{j \neq i}$ is a shorthand for: $\sum_{j=1, j \neq i}^{N}$.
[14] $\overline{x_i}$ denotes the mean of x_i over all i.

perceive firms to be either appealing or not. In reality, actors may perceive firms to be less or more appealing. To capture that intuition, the appeal of a firm to a resource should be represented by a continuous relation. We define p^*_{ir} to be a continuous appeal relation between a firm i and a resource r.

4.2.7. DEFINITION. Continuous appeal relation p^*
p^* is a relation between $i \in I$ and $r \in R$. So, $p^* \in P^* = I \times R$. For all $i \in I$ and $r \in R$ it holds that $0 \leq p^*_{ir} \leq 1$, dependent on the degree of appeal of i to r.

P^* is a continuous bimodal network of p^* relations.

4.2.9 From continuous appeal relations to competition

Intuitively, the intensity of competition in a pair of firms about a resource is dependent on the appeals that both firms make to the resource. We want to make that notion more formal, capturing the following intuitions: (1) competition between firms i and j about resource r is intense if i and j strongly appeal to r, (2) competition between i and j about r is 0 if either i or j does not appeal to r, and (3) competition between i and j about r is more intense if i and j make more similar appeals to r.

A straightforward way to capture intuition (1) and (2) simultaneously, is to let competitive intensity in pair i, j for r, which we call[15] v^*_{ijr}, be proportional to the *smallest* of the two appeals p^*_{ir} and p^*_{jr}, that is $min(p^*_{ir}, p^*_{jr})$.[16] Equation 4.4 expresses that notion.

$$v^*_{ijr} \propto min(p^*_{ir}, p^*_{jr}) \tag{4.4}$$

where $\propto$ indicates proportionality.

To capture intuition (3) we define *substitutability*. In our intuition, substitutability denotes the lack of preference for one firm over another. Note that firms do not necessarily have to be similar in order to be substitutable, for as long as they have similar appeals.

In order to define substitutability as a lack of preference, we first define preference. Preference, pr_{ijr}, for firm i over firm j, according to resource r, we define as the difference in the appeals of both firms. So:

4.2.8. DEFINITION. Preference pr_{ijr} for i over j according to r
$pr_{ijr} = p^*_{ir} - p^*_{jr}$

Note that $-1 \leq pr \leq 1$, and that pr is, in network terms, not symmetric. In fact, $pr_{ijr} = -pr_{jir}$.

If (the controller of) resource r perceives firm i as highly preferable over firm j, firm i will find it easy to employ r. As a consequence, the competitive intensity

[15] The superscript $.^*$ denotes that v^* is based on the p^*-relation, not on the p-relation.

[16] We calculate $min(x, y)$ as: $min(x, y) = \frac{(x+y)-|x-y|}{2}$. Note that $|z|$ here denotes the absolute value of z.

imposed by i and j's rivalry about r will be marginal. If r has no clear preference, and perceives the firms as substitutable, competitive intensity about r will be high. The definition of substitutability is:

4.2.9. Definition. Substitutability sub_{ijr} of i and j according to r
$sub_{ijr} = 1 - |pr_{ijr}|$

Notice that $0 \leq sub \leq 1$. Notice furthermore that sub is symmetric; $sub_{ijr} = sub_{jir}$.

We assume that:

$$v^*_{ijr} \propto sub_{ijr} \tag{4.5}$$

Equations 4.4 and 4.5 imply Definition 4.2.10.

4.2.10. Definition. Competitive intensity in pair i, j about r

$$v^*_{ijr} = min(p^*_{ir}, p^*_{jr}) \cdot sub_{ijr}$$

Notice that v^*_{ijr} can be re-written as follows:

$$v^*_{ijr} = \frac{(p^*_{ir} + p^*_{jr}) - |p^*_{ir} - p^*_{jr}|}{2} \cdot sub_{ijr} = \frac{((p^*_{ir} + p^*_{jr}) - |pr_{ijr}|) \cdot (1 - |pr_{ijr}|)}{2}$$

Figure 4.8 shows that v^*_{ijr} behaves according to our three intuitions. Notice that

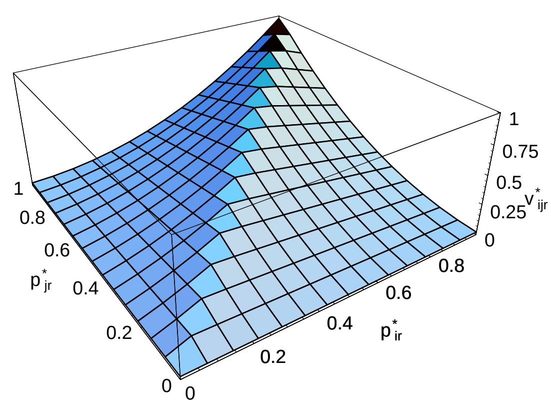


Figure 4.8: Competitive intensity as a function of two appeals

$0 \leq v^*_{ijr} \leq 1$ and that $v^*_{ijr} = v^*_{jir}$.

The total competition intensity v^*_{ij} in pair $i, j \in I$ is the aggregate of their competitive intensities v^*_{ijr} over all r.[17] So,

[17]By simply aggregating over the resources, we implicitly assume all resources to be of similar value. In reality, competition about more valuable resources may exert more pressure on the competing firms; resource values may differ, and firms may even perceive the value of resources differently. For an extension to our network-representation of competition that captures these intuitions, see (Vermeulen 2002a).

4.2.11. Definition. Competitive intensity in pair i, j, revisited

$$v^*_{ij} = \sum_{r=1}^{R} v^*_{ijr}$$

Notice that $0 \leq v^*_{ij} \leq R$, and that $v^*_{ij} = v^*_{ji}$.

The competitive pressure that v^*_{ijr} imposes on firm i is:

4.2.12. Definition. Competitive pressure on i by j, revisited

$$c^*_{ij} = \frac{v^*_{ij}}{v_i}$$

Notice that $0 \leq c^*_{ij} \leq 1$, and that c is asymmetric.

The total competitive pressure on i by its competitors is:

4.2.13. Definition. Competitive pressure on i, revisited

$$c^*_i = \sum_{j \neq i} c^*_{ij}$$

Notice, $0 \leq c^*_i \leq (N-1)$.

Finally, competitive pressure in the industry is:

4.2.14. Definition. Industry competition, revisited

$$C^* = \sum_{i=1}^{N} c^*_i$$

Notice, $0 \leq C^* \leq (N^2 - N)$. At this point, the micro-macro link is again established, also for markets with continuous appeal relations.

4.2.10 Niche size and the distribution of competitive pressure

As we have seen, not all firms in an industry need to face the same competitive pressure. In this section we will further investigate the role of niche size in competition. Specifically, we will try to determine the relation between the distribution of niche sizes and the distribution of competitive pressure over an industry.

In organization sociology, the competitive repercussions of size are a debated issue. On the one hand, Carroll (1985) and Carroll and Hannan (2000) assume smaller firms to be easily outcompeted if they engage in competition with larger firms. On the other hand, Barnett (1997) identified large firms as "weak competitors," which impose little danger on their smaller competitors.

In Section 4.2.6 we showed that niche size is an important determinant of competitive pressure. As an illustration of the degree of importance of niche size in competition, consider the ratio of competitive pressures c_{ij} and c_{ji}. For this ratio it holds that:

$$\frac{c_{ij}}{c_{ji}} = \frac{v_{ij}/v_i}{v_{ji}/v_j} = \frac{v_j}{v_i}$$

So, if the niche of i is n times larger than the niche of j, the competitive pressure on i by j, c_{ij}, is n times smaller than the competitive pressure on j by i.

If niche size is an advantage in a pairwise competitive relation, it could be a competitive advantage in industries too. To find out whether that conjecture is correct, consider the ratio of competitive pressure c_i and c_j. For this ratio it holds that:

$$\frac{c_i}{c_j} = \frac{\sum_{k\neq i} v_{ik}/v_i}{\sum_{l\neq j} v_{jl}/v_j} = \frac{\sum_{k\neq i} v_{ik}}{\sum_{l\neq j} v_{jl}} \cdot \frac{v_j}{v_i}$$

We estimate v_{ik} as follows: suppose that the niches of firms i and k, R_i and R_k, are located independently in resource base R. Then, the probability for each $r \in R$ to be included in R_i is $\frac{v_i}{|R|}$. So, the probability of each $r \in R$ to be included in both R_i and R_k is $\frac{v_i}{|R|}\frac{v_k}{|R|}$. The estimated size of v_{ik} is:

$$\hat{v}_{ik} = \frac{v_i}{|R|}\frac{v_k}{|R|}|R| = \frac{v_i v_k}{|R|}$$

As a consequence, c_i can be estimated as:

$$\hat{c}_i = \frac{\sum_{k\neq i} \hat{v}_{ik}}{v_i} = \frac{\sum_{k\neq i} v_i v_k}{v_i|R|} = \frac{\sum_{k\neq i} v_k}{|R|}$$

Let $R_N = \sum_{i=1}^{N} v_i$, then $\sum_{k\neq i} v_k = R_N - v_i$. Now, the ratio between c_i and c_j is estimated as:

$$\widehat{\left(\frac{c_i}{c_j}\right)} = \frac{\hat{c}_i}{\hat{c}_j} = \frac{\sum_{k\neq i} v_k}{\sum_{l\neq j} v_l}\frac{|R|}{|R|} = \frac{R_N - v_i}{R_N - v_j}$$

Usually, terms v_i and v_j are small compared to R_N, and then the ratio of c_i and c_j is close to 1: c_i and c_j are similar. Figure 4.9 shows the estimated distribution of competitive pressure in an industry where niche size is distributed according to a power law.[18] The figure shows that the distribution of competition over

[18]The assumption of niche sizes being distributed according to a power law is based on the idea that niche size is related to firm size, and that firm sizes are observed to be often distributed according to a power law (Ijiri and Simon 1977). Figure 4.9 is based on the following parameters: N=100, size of the r^{th}-greatest niche is given by $\frac{M}{r^{\beta}}$, where M, the size of the largest niche, is 1000, $\beta = 0.7$. $|R|$ is assumed to be 2000. Changing these values within the corresponding domains ($M > 0$, $N > 1$, $0 < \beta < 1$) does not change the characteristics of the figure.

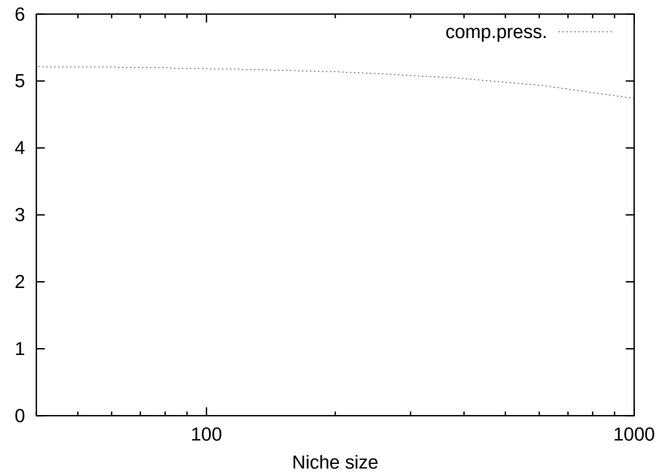


Figure 4.9: Distribution of competition over an industry

an industry can be expected to be approximately even. So, niche size is not a substantial competitive advantage in industries.

In contrast, niche size is an important variable in the competitive pressure *imposed by* firms. Let $c_{\bullet i}$ denote the total competitive pressure imposed by firm i.

$$c_{\bullet i} = \sum_{k \neq i} c_{ki} = \sum_{k \neq i} \frac{v_{ki}}{v_k}$$

v_{ki} is estimated as:

$$\hat{v}_{ki} = \frac{v_k v_i}{|R|}$$

So, $c_{\bullet j}$ is estimated as:

$$\hat{c}_{\bullet i} = \sum_{k \neq i} \frac{v_i v_k}{|R| v_k} = \sum_{k \neq i} \frac{v_i}{|R|} = \frac{(N-1) v_i}{|R|}$$

The ratio of $c_{\bullet i}$ and $c_{\bullet j}$ is estimated as:

$$\widehat{\left(\frac{c_{\bullet i}}{c_{\bullet j}}\right)} = \frac{\hat{c}_{\bullet i}}{\hat{c}_{\bullet j}} = \frac{(N-1) v_i / |R|}{(N-1) v_j / |R|} = \frac{v_i}{v_j}$$

So, the competitive pressure imposed by firms is estimated to be proportional to the size of their niche, whereas the competitive pressure received by firms is estimated to be (almost) unaffected by niche size.

This concludes our discussion of the relation between niche size and competitive pressure.

Hannan and Carroll, (1992, p.39) argue that in the ecology of organizations "[...]

competition depends on both the degree of intersection of fundamental niches and the number of competitors involved. It may also depend on the sizes of individual competitors [...]."

Our measures of niche intersection, v_{ij} and v^*_{ij}, and the measures of competitive pressure that are defined on the basis of it, capture and specify that intuition. The degree of niche intersection defines the intensity of competition between a pair of individual firms. The competitive intensity, together with the niche size of the respective competitors, determines the resulting competitive pressure. The number of firms in an industry, finally, determines the number of (potential) competitive relations, and thus the total pressure for the industry.

We stop modeling competition here. In the following section, we will apply our network representation of competition on an empirical domain: the market of Internet search engines.

4.3 Measuring competition: the market of Internet search engines

We start with some background information about the market of Internet search engines, and with formulating a number of claims relating to competition in the search engine market. Then we discuss the collection of data. Subsequently, we present the empirical data, and test the claims.

4.3.1 Internet search engines: background

Internet search engines are commercial businesses, providing indispensable Internet navigation tools. Together, search engines constitute an industry. Although users of search engines do not pay for the search service, they are principal resources. While using search engines, users look at advertisement banners, are referred to sites that pay to be referred to, have their search profiles analyzed, or are subject to other schemes that in one way or another add to the search engines' benefit. Although search engines also compete for advertisers, programmers, bandwidth, and technology, first and foremost they compete for users (Gandal 2001).

Search engines have been around for almost as long as the Internet; they can be traced back to 1993. From that time onward, the number of users of the Internet, and, consequently, of Internet search engines, has grown considerably. The number of search engines has also grown; in 2001, when we collected our data, we found traces of 137 more or less independently operating search engines that have been present at one time or another on the Net.

Over the years, the market strategies of newcomers in the search engine industry seem to have changed. Search engines that originate before 1997 mostly pursue a generalist strategy, by offering their services in English and by aiming at a world wide audience. Search engines that entered the market from 1997 onward, generally offer their services in a language other than English or to a particular geographical area (for a more comprehensive discussion of the search engine industry, see Chapter 5).

4.3.2 Competitive dynamics in the search engine industry

The changing competitive conditions in the search engine industry, as well as the changing competitive strategies of new search engines, can be expected to have an impact on the industry's competitive dynamics. First, the growing number of search engine industry members should increase industry competition. Second, the specialist strategy of most later entrants should reduce the competitive pressure imposed on them; as a consequence, the mean firm-on-firm competitive pressure in the industry should eventually decrease, which, in turn, reduces the aggregate industry competition. Third, the increasing number of Internet users increases the probability that—even by random chance—two search engines compete for a single user; consequently, the density of the competitive network can be expected to increase.

Propositions 4.3.1 to 4.3.3 formulate these expectations:

4.3.1. PROPOSITION. *N increases.*

4.3.2. PROPOSITION. *$\overline{c}_{ij}$ eventually decreases.*

4.3.3. PROPOSITION. *ΔS increases.*

If N increases (Proposition 4.3.1), then $(N^2 - N)$ increases at an increasing rate. By Equation 4.2, if ΔS increases, CR increases relatively to $(N^2 - N)$. So, as a corollary of Propositions 4.3.1 and 4.3.3, Proposition 4.3.4 states that CR increases at an increasing rate.

4.3.4. PROPOSITION. *CR increases at an increasing rate.*

Propositions 4.3.2 and 4.3.4 together imply that industry competition first increases at an increasing rate, but eventually (as $\overline{c}_{ij}$ decreases) at a decreasing rate. This will result in an S-shaped trajectory for industry competition.

4.3.5. PROPOSITION. *C increases first at an increasing rate, and later at a decreasing rate.*

Proposition 4.3.5 contrasts to organizational ecology theory, which predicts a exponential growth of industry competition. In Chapter 5 we discuss this issue at length.

Propositions 4.3.2 to 4.3.5 focused on the intensity of competitive pressure. The following two propositions focus on the distribution of competitive pressure.

According to the argument in Section 4.2.10, we expect competitive pressure to be distributed approximately proportional over firms with niches of different size. Proposition 4.3.6 formulates this expectation slightly differently, by stating that niche size and competitive pressure received are mutually independent.

4.3.6. PROPOSITION. *c_i and v_i are independent.*

Proposition 4.3.6 formulates the argument in Section 4.2.10, by stating that competitive pressure imposed by firms is positively related to niche size.

4.3.7. PROPOSITION. *$c_{\bullet i}$ and v_i are positively related.*

This concludes our propositions regarding the competitive dynamics in the search engine industry. We proceed by studying the empirical domain.

4.3.3 Measurements

Internet resources (*e.g.* web pages) are identified by so-called "Uniform Resource Locators" (URLs), also referred to as Internet addresses. All over the Internet, Internet resources mention URLs of other Internet resources; the so-called Internet "links." The fact that Internet resources link to each other gives the Internet its network structure.

Maurer and Huberman (2000) suggest that the number of Internet resources that mention a specific resource is indicative for the number of users of the latter resource. We follow that suggestion, and take the number of Internet pages that mention a specific Internet search engine to be indicative for the number of users of that search engine. Because users largely constitute the niches of search engines, an indication of the number of users of a search engine is an indication of the size of its niche. So, to determine search engine niche size, we measured, from October 1993 to October 2000, for 137 search engines:

- The number of Internet pages that mention the URL of a particular search engine, per yearly quarter.

Internet resources may mention more than one search engine URL. So, Internet pages may be in more than one search engine niche. In that case, there is niche intersection. To determine the degree of niche intersection, we measured, again from October 1993 to October 2000:

- The number of Internet pages that mention the URLs of a particular *pair* of search engines, per yearly quarter.

We collected the data using search engine Northernlight.com, which has a number of advantages over other engines.[19] We collected all data in early 2001, in

[19]For a discussion of the choice for Northernlight.com as the principle source for data, see Chapter 5.

a 10 day session. The "historical" character of the data is obtained by restricting the search to Internet pages that themselves were dated, and by ordering the pages found by the time they were last updated. For a more detailed account of the data, as well as about the methods used to collect the data, we refer to Chapter 5.

4.3.4 Data

Let $\mathbf{v}_{it}$ be the number of Internet pages mentioning the URL of search engine i in quarter t. Using the vocabulary of the previous chapter, $\mathbf{v}_{it}$ is equal to the number of relations p_{ir} between firm (search engine) i and resources (pages) r, at time t. If 10 or less pages that mention a particular URL are found, we say $\mathbf{v}_{it} = 0$. This threshold prevents small numbers of falsely dated Internet pages to obscure our measurements. So, $\mathbf{v}_{it} > 10$.

Let $\mathbf{v}_{ijt}$ be the number of Internet pages that mention the URLs of search engine i and search engine j at t. $\mathbf{v}_{ijt}$ has no threshold, so $\mathbf{v}_{ijt} \geq 0$.

Whereas $\mathbf{v}_{it}$ is an indicator for v_i (the measure of niche size), $\mathbf{v}_{ijt}$ is an indicator for v_{ij} (the measure for competitive intensity). In the previous section, we calculated c_i on the basis of v_i and v_{ij}. Similarly, we now calculate $\mathbf{c}_{it}$ on the basis of $\mathbf{v}_{it}$ and $\mathbf{v}_{ijt}$. The measure for industry competition at time t, $\mathbf{C}_t$, is obtained by aggregation of $\mathbf{c}_{it}$ over all i.

In addition, we operationalize $\mathbf{s}_{ijt}$, an indicator for s_{ij}, the dichotomous measure of presence of a competitive relation, as follows:

1. Operational definition.
$\mathbf{s}_{ijt} = 1$ *iff* $\mathbf{v}_{it} > 10$, $\mathbf{v}_{jt} > 10$ *and* $\mathbf{v}_{ijt} > 0$; $\mathbf{s}_{ijt} = 0$ *otherwise.*

We obtain $\mathbf{CR}_t$ from $\mathbf{s}_{ijt}$ by means of aggregation.

4.3.5 Empirical results

The data collected in the industry of Internet search engines is displayed in the table in Appendix I. Let us review, one by one, our propositions regarding the competitive dynamics in the search engine industry in this period.

Proposition 4.3.1:
Proposition 4.3.1 states that the number of search engines, N, increases. Figure 4.10 shows the number of search engines per yearly quarter, from 1993 to 2000. It shows that in this period, the Internet search engine industry has grown from scratch to more than 120 members.

Proposition 4.3.2:
Proposition 4.3.2 states that $\bar{c}_{ij}$ eventually decreases, due to the changing competitive strategies of new entrants in the search engine industry. Figure 4.11

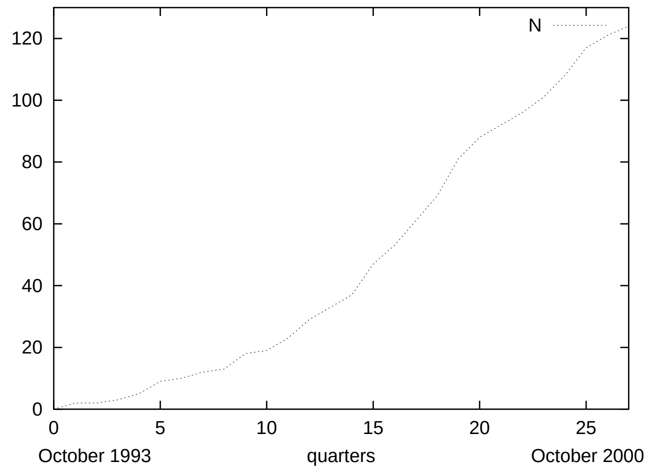


Figure 4.10: Number of firms in Internet search engine market

confirms this proposition. It shows that, from approximately t_{21} onward,[20] $\overline{c}_{ij}$ decreases.

Proposition 4.3.3:
Proposition 4.3.3 states that the density of the S-network, that is, the ratio of actual competitive relations to potential competitive relations, increases, due to the growth of the resource base. Figure 4.12 confirms this proposition. It shows that, from approximately t_{13} onward,[21] the density of the S-network increases.

Proposition 4.3.4:
Proposition 4.3.4 states that, as a consequence of the increase of both the size of the industry, N, and the density of the competition network, ΔS, the number of competitive relations in the industry, CR_t, increases at an increasing rate. This proposition is confirmed by Figure 4.13.

Proposition 4.3.5:
Proposition 4.3.5 states that, as a consequence of Propositions 4.3.2 and 4.3.4, industry competition, C, has an S-shaped growth trajectory. Figure 4.13 confirms this proposition; the S-shaped pattern is clearly visible. Until t_{21} (January 1999), industry competition increases at an increasing rate. From approximately t_{21} onward, C decelerates.

Appendix II displays correlations between niche size, on the one hand, and competitive pressure, on the other. It also displays the correlations between niche-size

[20] t_{21} denotes January 1999.
[21] t_{13} denotes January 1997.

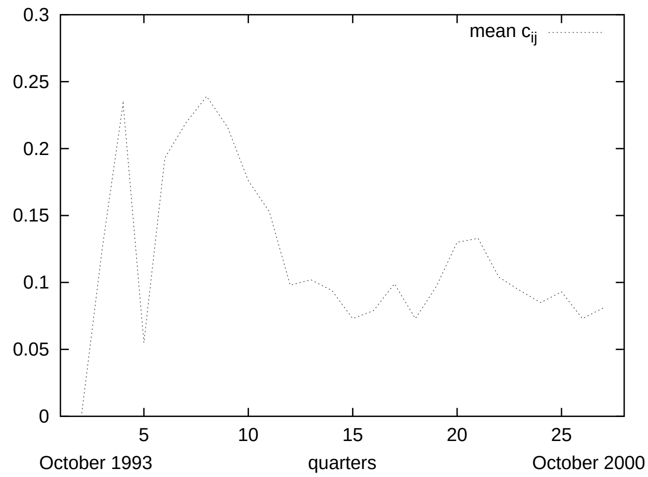


Figure 4.11: Mean firm-on-firm competitive pressure $\overline{c}_{ij}$.

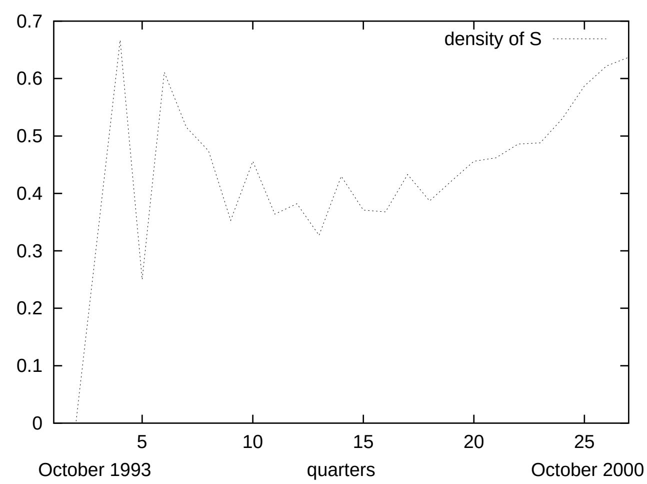


Figure 4.12: Density of the competition network ΔS.

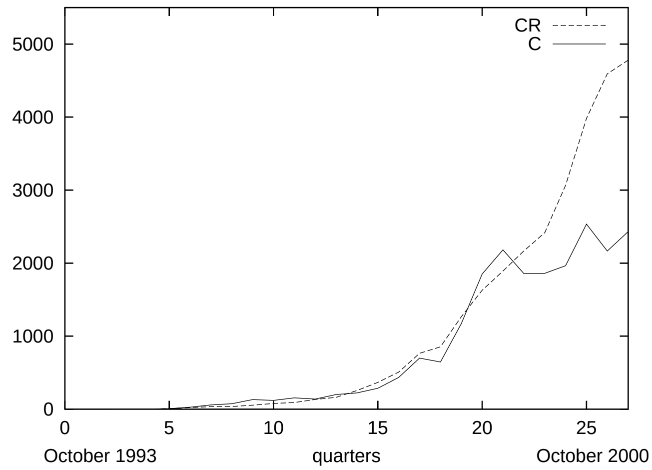


Figure 4.13: Number of competitive relations CR, and industry competition C.

rank and competitive pressure. Niche-size rank is the rank of a firm in a niche size ordering. The firm with the largest niche has rank 1 and the firm with the smallest niche has rank N. The niche-size rank of firm i, we call r_i.

To be precise, Appendix II displays: (1) the correlation between niche-size rank, r_i, and competitive pressure that firms receives from all other firms, c_i, (2) the correlation between niche size, v_i, and the received competitive pressure c_i, (3) the correlation between niche-size rank and the competitive pressure that i imposes on all other firms, $c_{\bullet i}$, and (4) the correlation between niche size and the imposed competitive pressure. We use these correlations to test Propositions 5 and 6. We only take into account correlations if $N > 10$.

Proposition 4.3.6:
Proposition 5 states that the competitive pressure received by a firm, and its niche size, are independent variables. Appendix II displays, for all quarters in our focal period, low correlations between niche-size rank and competition received.[22] Most correlations are positive, which indicates that large firms—with low ranks—receive slightly less competitive pressure than do small firms. The correlations between niche size and received competitive pressure show, by and large, the same result. All correlations are low. Moreover, most correlations are negative, which again indicates that large firms receive slightly less competitive pressure than do small firms. The low correlations lead us to conclude that received competitive pressure and niche size are independent, which confirms Proposition 4.3.6.

As an illustration, Figure 4.14 shows the competitive pressure received by the search engines in the industry, at t_{24} (the last quarter of 1999). Note that the x-

[22] Although for 5 out of the 21 quarters the correlations are significant, all correlations are below 0.3.

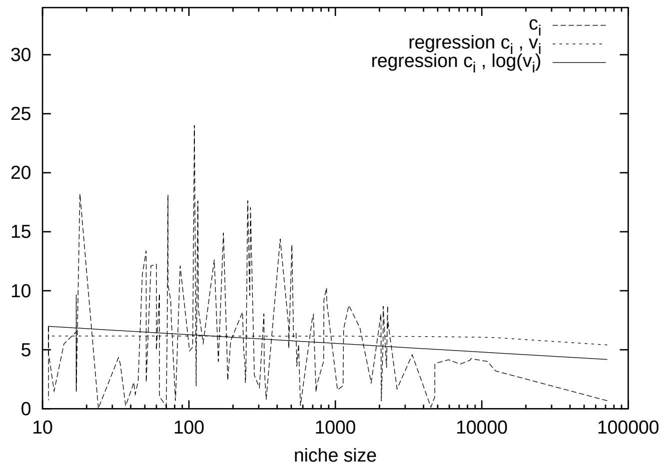


Figure 4.14: Competitive pressure received, with niche size, at t_{24}

axis, representing niche size v_i is in logarithmic scale. The figure shows that firms receive varying degrees of competitive pressure. To elucidate possible trends, we have added the optimal linear and logarithmic regression lines.[23] The lines show that at t_{24}, large firms receive approximately the same competitive pressure as small firms do.[24]

Proposition 4.3.7:
Proposition 4.3.7 states that the competitive pressure imposed by a firm is positively related to its niche size. In Appendix II shows the correlations between niche-size rank and imposed competition for all quarters in our focal period. All correlations are strongly negative and significant. So, large firms—with low ranks—impose more competitive pressure on their peers than do small firms. The correlations between niche size and imposed competition show corresponding results. All correlations are high, and all are significant: large firms, again, impose more competitive pressure than do small firms. The high correlation values lead us to conclude that the niche size and competitive pressure imposed om peers, are positively related, which confirms Proposition 4.3.7.

As an illustration, Figure 4.15 shows the competitive pressure that firms with different niche sizes impose on their peers, at t_{24}, the quarter that was also represented by Figure 4.14. Again, niche size v_i is represented in logarithmic scale.

[23]The optimal linear regression is: $\hat{c}_{it} = 6.16 - 0.00001 \cdot v_{it}$. For this regression we have $R^2 = 0.024$ and $T = -1.63$. The optimal logarithmic regression is $\hat{c}_{it} = 7.75 - 0.32 \cdot log(v_{it})$. For this regression, $R^2 = 0.015$ and $T = -1.28$. The slopes of both regressions are mild; moreover, the regressions explain only a limited degree of the variance.

[24]Note, that because of the logarithmic scale of the x-axis, Figure 4.14 displays the linear regression as a curve, and the logarithmic regression as a straight line.

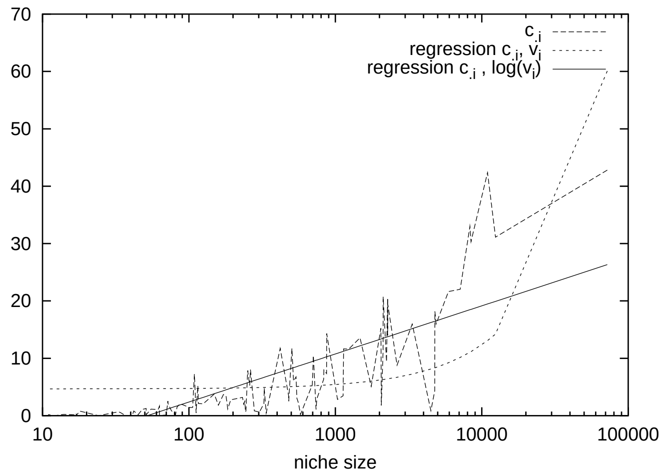


Figure 4.15: Competitive pressure imposed on peers, with niche size, at t_{24}.

Figure 4.15 shows that firms impose different degrees of competitive pressure on their peers. To elucidate possible trends, we have again added the optimal linear and logarithmic regression lines.[25] The lines show a clear trend: large firms impose more competitive pressure than do small firms.

To sum up the empirical results, competitive pressure in the industry of Internet search engines increases. This is not due to an increase of the firm-on-firm competitive pressure—in fact, firm-on-firm competitive pressure rather decreases than increases—but due to (1) the increasing number of competitors and (2) an increasing density of the competition network. Furthermore, firms with large niches impose more competitive pressure than firms with small niches, but receive by and large the same competitive pressure.

4.4 Summary and results

The goal of this chapter was to develop a network-representation of competition that related competitive processes at the level of individual firms to competitive processes at the industry level. The network was intended to generate quantifiable measures for a number of basic notions of competition. The variables in the network representation were intended to have straightforward operational definitions, making the notions of competition directly measurable.

We started out by making a distinction between three entities that all are

[25] The optimal linear regression is: $\hat{c}_{\bullet it} = 4.67 + 0.0008 \cdot v_{it}$. For this regression it holds that $R^2 = 0.40$ and $T = 8.37$. The optimal logarithmic regression is $\hat{c}_{\bullet it} = -14.38 + 3.64 \cdot log(v_{it})$. For this regression, $R^2 = 0.61$ and $T = 12.89$.

called "competition" in the literature. First, the competitive relation between firms; second, the intensity of this competitive relation; and third, the competitive pressure that the relation imposes on both competitors.

We defined a competitive relation to be present if two firms appeal to the same resource. The intensity of the competitive relation is determined by the number of resources to which two firms make a coinciding appeal. The competitive pressure imposed by the relation is determined by the proportion of a firm's niche to which the other firm appeals. If firms have niches of different sizes, competitive pressure is not symmetric.

Having made these distinctions, relating competitive processes at firm and industry level turned out to be straightforward. The competitive relations between all members of an industry together constitute a competition network. By establishing the size and the density of the network, we obtain the number of competitive relations in an industry.

The intensity of the competition between individual firms can be represented as a continuous relation. Thus, a network of competitive intensities can be represented as a network of continuous relations. This network contains more information about how firms relate to each other, than does the network of dichotomous competitive relations.

Competitive pressure, as said, can be asymmetric in a pair of firms. Consequently, a network of competitive pressures is not only continuous, but also directed. By aggregating the competitive pressure exerted by all firms in an industry on a focal firm, we obtain the total competitive pressure that the focal firm faces. By aggregating the competitive pressure on all firms in an industry, we obtain the total competitive pressure on the industry. The latter measure falls within the definition of industry competition as stated by organizational ecology. So, by means of aggregation, competition at the individual level, and at the industry level, are formally related.

In Section 4.2.8, we returned to the micro level of analysis and implemented the notion that appeals of firms to resources are better represented by continuous than by dichotomous relations. We defined the continuous relations, and used them to define the notions of preference and substitutability, on which, in turn, we defined an alternative notion of competitive intensity.

In the next section we showed that niche size is an important factor in a pairwise competitive relation. In a pair of firms, the firm with the larger niche imposes more competitive pressure, and receives less competitive pressure. If we consider competition in industries, the competitive advantage of having a large niche for the most part disappears. Firms with large niches impose more competitive pressure than do firms with small niches, but the pressure that large firms receive is by and large the same.

In Section 4.3, we showed that the presence of competitive relations, their intensity, and the pressure they exert, can be directly measured. We studied the industry of Internet search engines, and assumed that the number of Internet pages that mention the URL of a particular search engine is representative for the number of resources of the search engine. We counted the number of pages that mentioned the URLs of particular search engines individually, and those that mentioned the URLs of particular pairs of search engines; the latter count measured niche intersection. By collecting the data per quarter, over a period of seven years, we were able to determine the competitive dynamics in the Internet search engine industry. The data showed an increase of industry competition according to an S-shaped pattern. This pattern arose as a result of, on the one hand, an increasing size and density of the competition network, and, on the other hand, a decreasing pressure imposed by the competitive relations.

4.5 Discussion and conclusions

In this chapter, we analyzed competition by means of a network-representation. The basic representation is a bimodal market network. Upon the appeal relations that constitute the market network, we define competitive relations between firms. Although we considered both dichotomous as continuous appeals, we did not analyze the appeal-relations themselves, and focused on the competitive relations between the firms instead. Together, the relations between the firms constitute a competition network.

One reason for applying network analysis to competition is that network analysis has a mathematical apparatus in which several properties that characterize a network are readily formalized. In the current chapter, the main interest was in simple measures, such as the size and the density of the competitive network, but for other purposes other measures may prove interesting. Examples are: (1) the diameter of the network (the number of steps lying between the nodes that are the furthest removed); (2) its connectivity (the number of ways to go from one node to another); (3) the presence of a core/periphery structure (here: the presence of a center of competition) (Borgatti and Everett 1999); and (4) the degrees of the network's nodes (the number of other nodes to which nodes are connected). Investigating an industry by analyzing its competition network makes it possible to formally characterize the industry; but also to mutually compare industries and investigate the structural differences between them, and to investigate their competitive dynamics by use of evolving network analysis (Vermeulen 2002b).[26]

A particularly interesting formal property of competition networks is their dimensionality (Steyvers 2002). For one, analyzing dimensionality provides insights in the considerations of the consumers in choosing one firm's product over another's; an insight that can be used for marketing purposes. Moreover, dimensionality is hypothesized to be an important factor in the viability of new entrants

[26]Possibly, such an investigation could elucidate the phenomenon of preferential attachment of social actors, which has been widely discussed, but rarely observed (Barabási and Albert 1999).

in the market (Péli and Nooteboom 1999).

Another advantage of studying competitive relations by means of network analysis, is that the appropriate data may be easily obtained from the Internet. The procedure through which we collected data about the Internet search engine market, can be duplicated for other industries, especially when—in the near future—other industries may become more visible on the Net.

In our empirical set up, the measurement used to indicate competition is the co-reference to search engines on Internet pages. One might wonder whether co-occurring references are really indicative for competition. If an Internet page mentions two search engines, does that necessarily mean the author of the page considers the search engines to be substitutable, as is our implicit assumption? Maybe the Internet author considers the search engines to be complementary, which would implicate that we mistake competition for some sort of co-habitation of firms.

For two reasons we believe, however, that a co-reference to firms implies substitutability, and thus competition, rather than co-habitation. First, we observed competitive relations between search engines that operate in the same language or geographical area, to be the most intense (see also Chapter 5). Second, co-habitation of firms implies that the use of one search engine instigates the use of another search engine. This contradicts the whole purpose of search engines: intended to answer queries, not to evoke new queries.

We have stated that our notion of competitive intensity formalizes the organizational ecological notion of niche intersection. In (Sohn 2001) a number of criteria are defined that a measure for organizational niche intersection should meet. First, the measure should be 0 if there is no intersection and 1 if there is intersection; second, it should be in ratio scale. Third, it should take size differences into account. Our measure of niche intersection meets all three criteria.

Nevertheless, there are two important differences between Sohn's measure of niche intersection and ours. First, Sohn assumes niche intersection and competition in a pair of firms to be identical. We showed that a competitive relation can impose different pressures on the firms involved. Second, Sohn estimates niche intersection from the dissimilarities between the individual resources in the niches, in the tradition of McPherson (1983), Levins (1968), and Hannan and Freeman (1977), who determine the Euclidean distance between so-called "competition coefficients," in order to estimate niche intersection. In contrast, we have measured niche intersection directly. As a consequence, we did not influence our measurements by first interpreting resource dissimilarities. In Chapter 5 of this thesis, we will look at the relation between niche intersection and resource (dis)similarity from a different angle, and define a measure of "niche similarity."

Some words should be spent on the definition of niche that we applied in this chapter. We defined the niche of a firm as a set of resources; the size of the niche is the number of resources for which a firms competes.

Admittedly, organizational ecology employs a more complex notion of niche, which, moreover, applies to industries, not to individual firms. The original organizational ecological definition of a niche was as follows: "The niche, then, consists of all those combinations of resource levels at which the population can survive and can reproduce itself" (Hannan and Freeman 1977, p.947). A set consisting of "combinations of resource levels" is a complicated mathematical object. The intersection of two sets of combinations of resource levels is even more complex. As it turns out, no attempts have ever been made to measure niche intersection according to the ecological definition (Sohn 2001); in general, members of the same industry were simply assumed to have the same niche (Hannan and Carroll 1992, p.29).

By defining niches as sets of resources, we may have oversimplified matters. However, it was precisely this simplification that allowed us to determine niche intersection, and to define competitive relations, their intensity, and the pressure they impose on firms. Our simplification also allowed us to establish the micro-macro link in competition. And last but not least, it allowed us to measure competition directly; not by merely counting the number of firms in an industry; not by making empirical generalizations about competitive advantages that one type of firm might have over another type; not by imposing our own perception of resource dissimilarities; but by measuring the appeals that search engines make to their resources: the Internet users.

In sum, we specified a network-representation of competition in industries, thereby formally connecting competition at the (micro) level of individual firms, and competition at the (macro) level of industries. We measured the intensity of competitive relations in pairs of Internet search engines, and the competitive pressure imposed on all search engines on the Internet. By use of the micro-macro link in competition, we were also able to the measure the competitive dynamics of the search engine industry.

Appendix I

The table below shows for each quarter t, the number of search engines operating on the market, N_t; the potential number of competitive relations, $\frac{N_t^2 - N_t}{2}$; the actual number of competitive relations, CR_t; the industry competition, C_t; the density of the competitive network, ΔS_t; and the mean firm-on-firm competitive pressure, $\overline{c}_{ijt}$.

Year	Quarter	$\mathbf{N}_t$	$\frac{\mathbf{N}_t^2-\mathbf{N}_t}{2}$	$\mathbf{CR}_t$	$\mathbf{C}_t$	$\Delta\mathbf{S}_t$	$\overline{\mathbf{c}}_{ijt}$
1993	4th	0	0	0	0.00	undef.	undef.
1994	1st	2	1	0	0.00	0.000	0.0000
	2nd	2	1	0	0.00	0.000	0.0000
	3rd	3	3	1	0.15	0.333	0.0125
	4th	3	3	2	0.28	0.667	0.0234
1995	1st	9	36	9	0.79	0.250	0.0055
	2nd	9	36	22	2.77	0.611	0.0193
	3rd	12	66	34	5.77	0.515	0.0219
	4th	13	78	37	7.45	0.474	0.0239
1996	1st	18	153	54	13.22	0.353	0.0216
	2nd	19	171	78	12.04	0.456	0.0176
	3rd	23	253	92	15.53	0.364	0.0153
	4th	27	351	134	13.82	0.382	0.0098
1997	1st	32	496	162	20.16	0.327	0.0102
	2nd	35	595	256	22.30	0.430	0.0094
	3rd	45	990	367	28.73	0.371	0.0073
	4th	53	1378	507	43.51	0.368	0.0079
1998	1st	60	1770	766	70.01	0.433	0.0099
	2nd	67	2211	855	64.72	0.387	0.0073
	3rd	78	3003	1266	116.98	0.422	0.0097
	4th	85	3570	1629	185.07	0.456	0.0130
1999	1st	91	4095	1892	218.21	0.462	0.0133
	2nd	95	4465	2168	185.78	0.486	0.0104
	3rd	100	4950	2417	186.12	0.488	0.0094
	4th	108	5778	3065	196.41	0.530	0.0085
2000	1st	117	6786	3982	253.48	0.587	0.0093
	2nd	122	7381	4593	216.64	0.622	0.0073
	3rd	123	7503	4780	243.29	0.637	0.0081

Appendix II

The table below shows for each quarter t, the number of search engines operating on the market, N_t; the (Pearson) correlation between niche-size rank, r_{it}, and the competitive pressure received by a firm, c_{it}; the correlation between r_{it} and the competitive pressure imposed by a firm, $c_{\bullet it}$; the correlation between niche size., v_{it}, and the competitive pressure received, c_{it}; and the correlation between v_{it} and the competitive pressure imposed, $c_{\bullet it}$. * denotes a two-tailed significance $P \leq 0.05$. ** denotes a two-tailed significance $P \leq 0.01$.

Year	Quarter	$\mathbf{N}_t$	$corr(r_{it}, c_{it})$	$corr(r_{it}, c_{\bullet it})$	$corr(v_{it}, c_{it})$	$corr(v_{it}, c_{\bullet it})$
1993	4th	0	undef.	undef.	undef.	undef.
1994	1st	2	undef.	undef.	undef.	undef.
	2nd	2	undef.	undef.	undef.	undef.
	3rd	3	0.99	0.59	-1.00	-0.70
	4th	3	0.98	0.98	-0.84	-0.84
1995	1st	9	0.24	-0.33	0.00	0.53
	2nd	9	-0.31	-0.78**	0.29	0.77**
	3rd	12	-0.08	-0.83**	0.03	0.92**
	4th	13	0.25	-0.75**	-0.13	0.85**
1996	1st	18	-0.16	-0.81**	0.04	0.87**
	2nd	19	0.05	-0.74**	-0.13	0.69**
	3rd	23	-0.22	-0.76**	0.05	0.73**
	4th	27	-0.23	-0.79**	0.06	0.85**
1997	1st	32	-0.26	-0.75**	0.06	0.86**
	2nd	35	-0.17	-0.76**	0.07	0.87**
	3rd	45	0.17	-0.70**	-0.20	0.83**
	4th	53	0.24	-0.65**	-0.26	0.74**
1998	1st	60	0.18	-0.71**	-0.20	0.83**
	2nd	67	0.17	-0.68**	-0.18	0.67**
	3rd	78	0.24*	-0.68**	-0.18	0.64**
	4th	85	0.23*	-0.70**	-0.18	0.65**
1999	1st	91	0.11	-0.74**	-0.16	0.66**
	2nd	95	0.08	-0.73**	-0.15	0.61**
	3rd	100	0.10	-0.72**	-0.16	0.61**
	4th	108	0.10	-0.72**	-0.16	0.62**
2000	1st	117	0.27**	-0.74**	-0.17	0.61**
	2nd	122	0.24**	-0.70**	-0.17	0.62**
	3rd	123	0.20*	-0.70**	-0.17	0.63**

$* = P \leq 0.05$
$** = P \leq 0.01$

Chapter 5

Distribution of Competition in The Market of Internet Search Engines

5.1 Introduction

[1] This chapter investigates the emergence of the market of Internet search engines from the point of view of organizational ecology. It describes and explains the particular dynamics in the search engine market by the use of ideas and concepts from two organizational ecological theories: density dependence (Carroll and Hannan 2000, pp.213-228) and resource partitioning (Carroll and Hannan 2000, p.261-269). Although both theories share their parentage, as well as some basic assumptions, they differ in their account of the dynamics of the search engine market. The concept of competition plays a key role in both theories, but is defined differently. We will show that, through the conceptual refinement of competition, we can connect the two theories, and consequently, better understand and relate the phenomena observed in the market of Internet search engines.

The search engine market has a number of notable characteristics. First, obviously, it is a new market. The first search engines appeared in 1993 (Sonnenreich and Macinta 1998). Since then, the population of search engines has shown an impressive growth. At the time of data collection in 2001, around 150 engines operated independently, in all corners of the Internet.[2] Other features that make the market of search engines stand out are the explosive growth of its consumer base,[3] its next–to–perfect scale economies, its low barriers-to-entry (Gandal 2001), its lack of geographical constraints, its receptiveness to technological innovation, and the relative absence of price–competition (Shapiro and Varian 1999).

[1]Empirical data used in this chapter was courtesy of Northern Light Search. Equation 5.3 was suggested by Rob Mokken. Figures 5.2, 5.22, and 5.23 were created by Jaap Kamps.

[2]Source: `http://www.searchenginewatch.com`.

[3]Source: `http://www.nielsen-netratings.com`.

From an organizational point of view, the Internet seems to be an environment providing opportunities for search engines of every age, size and origin. Older search engines, on the one hand, have rode the wave of Internet's explosive growth, and, now that they have grown large, can reap scale economies. New, small organizations, on the other hand, benefit from the Net's lack of entry barriers, and from its openness to change and innovation. Moreover, entrants are not restrained by geography.

The search engines that populate the Internet today, constitute a large and differentiated group, offering services in many languages, and targeting many geographical domains and subject areas. One could wonder, though, whether the differentiation observed in the market of search engines is indeed caused by opportunities, or by necessities, instead. On the one hand, the Internet may be the lenient environment portrayed above; an environment allowing for denizens of all sorts and conditions. On the other hand, later entrants in the search engine market may have been forced to differentiate in order to avoid having to compete with the traditional, dominant players. In the latter case, the search engine market would be heading toward segregation: a "mass" market, controlled by a few large firms, and a specialist market, populated by smaller organizations.

Such segregation processes have been described by resource partitioning theory, a part of organizational ecology (Carroll 1985). Resource partitioning theory claims that segregation between a generalist and a specialist market is due to market concentration. The argument is that concentration causes large firms to move toward the market center, creating empty niche pockets in the market periphery for small new firms to enter into. Consequently, concentration causes the industry's resources to become partitioned.

Moreover, according to resource partitioning theory, concentration causes life chances for small specialist firms to improve. This last conjecture contradicts classical economics, which claims that market concentration creates barriers to entry for new organizations, and consequently has a negative effect on processes of differentiation that new entrants are assumed to entail (Schmalensee and Willig 1989).

Although it would be interesting to apply resource partitioning theory to the market of Internet search engines, on a first review such an exercise appears hobbled by the fact that the theory is formulated for mature industries, rather than for emerging ones. The argument of resource partitioning theory relies on the assumption that an industry is in an equilibrium state, where founding and disbanding rates balance each other out (see Chapter 3). In the period in which we investigated the industry of Internet search engines, founding rates dominated disbanding rates by far. We observed an industry that emerged, and was becoming differentiated, at the same time.

The part of organizational ecology that explains how markets emerge is density dependence theory (Hannan and Carroll 1992). This theory claims that founding and mortality rates are governed by legitimation and competition, which, in turn are governed by the number of firms in a population, the density.[4] Density dependence theory assumes all firms in an industry to target the same resources (*e.g.*, consumers), and consequently, to engage in full competition with each other. The theory does not accommodate for industries that differentiate while they emerge.

In this chapter, we aim to connect the theories of resource partitioning and density dependence. The idea of connecting both theories is not new (Hannan and Carroll 1992, p.48), but has never been realized. By connecting the theories we hope to be able to clarify the dynamics of organizational differentiation in emerging industries.

A factor that may complicate connecting both theories, is that resource partitioning is a descriptive theory, whereas density dependence theory is basically a mathematical model. In Chapter 3, we formalized the descriptive argument of resource partitioning theory in first-order logic. By use of that formalization, the implicative relations between the different "events" taking place are formalized; still, the theory's variables and their mutual relations are not mathematically specified. To facilitate a connection between resource partitioning and density dependence theory, we will in the proceeding sections formulate mathematical specifications of resource partitioning's key variables.

Subsequently, both organizational ecological theories can be connected. The key to connecting both theories is the measure of competition that we developed in Chapter 4. This measure connects the micro level of competition (competition between individual firms) to competition at the macro level (competition as a property of industries). The result is a notion of competition as a pressure that is distributed unevenly over industrial populations. The notion that competitive pressure is unevenly distributed over industries is fundamental to resource partitioning theory, and is not applied in density dependence theory. Therefore, resource partitioning and density dependence are complementary theories: whereas density dependence theory focuses on the intensity of competition, resource partitioning focuses on the distribution of competition.

By use of our mathematical model, we aim to answer the question: What are the dynamics governing the simultaneous emergence of, and differentiation in, the emerging Internet search engine industry?

Our question is not motivated solely by our observation of a seemingly differentiating population in the Internet search engine market. It is also motivated by a general interest in processes of differentiation in industries. Differentiation facilitates the entrance of new firms, which introduce new ideas and new standards in a market (Mezias and Mezias 2000). Thus, differentiation may stimulate innovation and raise product or service quality. Moreover, a differentiated population

[4]We will not use the term "density" to denote population size in this chapter; in Chapter 4 of this thesis, density is attributed a different meaning.

of suppliers allows consumers to choose, and thus may raise the level of consumer satisfaction.

This chapter is organized as follows. First, in Section 5.2, we will give an introduction in the Internet search engine industry, by means of a short historical account of the industry in the focal period of our empirical study—from 1993 to 2000. Next, in Section 5.3 we will briefly introduce the main ideas of organizational ecology, as well as its "sub-theories", resource partitioning, and density dependence. In this section, too, we show how the competition measure from Chapter 4 can be introduced in the mathematical model of density dependence. We furthermore show how legitimation of an organizational form, which plays an important role in density dependence theory, can be estimated from the propagation of a form's notoriety through a consumer population. In Section 5.4, we connect resource partitioning and density dependence, generating a theory that can explain resource partitioning processes in changing competitive environments.

In Section 5.5 we discuss the data we collected, which involves the measurement of pairwise competition between 137 individual search engines, in 28 quarterly assessments, from 1993 to 2000. We also connect the measurements to the theoretical concepts by means of operationalization. In Section 5.6, we discuss the results of the empirical study. We test our assumptions, and thus investigate both density dependence and resource partitioning, as well as their interplay.

In Section 5.7, we investigate two variables that resource partitioning theory claims to affect specialist life chances: market concentration and an increasing "dimensionality" of the market. At concentration we only look from an empirical angle; in order to measure market dimensionality, we have to resolve some theoretical issues first.

We conclude this chapter with a discussion.

5.2 The market of Internet search engines: a short history

In 1992, the first web browser was invented, and the internet as we now know it came into existence.[5] Early Internet users in search of information used Wandex, an index of all available web pages. As the Internet grew further, however, Wandex lost track of the information available on the Internet. At the same time, the program that Wandex used to collect its information began to significantly slow down Internet traffic (Sonnenreich and Macinta 1998). Consequently, the need arose for new tools that would allow users to search for information on the Internet. These came in two basic shapes: Internet directories, early examples of

[5] At `http://www.w3.org/History/19921103-hypertext/hypertext/WWW/Summary.html` the original summary of the WWW–project can still be found, as well as links to the first browser software.

which are ALIWEB and Yahoo, and Internet search engines, the first of which were JumpStation, the World Wide Web Worm (WWWW), and the Repository-Based Software Engineering Spider (RBSE), which are now all obsolete.

Directories and search engines were different not only in the way in which they categorized information, but also in how they extended their database. While directories relied on the users to "manually" provide and categorize new entries, search engines employed programs that automatically visited web–pages, and categorized the information on them in an index (Lawrence and Giles 1998).

The first search engines were usually pet projects of computer scientists,[6] but from 1994 onward professional search services such as Excite, Lycos, and AltaVista entered the market. From that time on the population of search engines started to grow at a steady pace. Currently more than a hundred search engines are in operation.[7]

5.2.1 Professionalism and differentiation

As search engines became more professional, they became more commercial. This was illustrated by the appearance of TV commercials, stock emissions, mergers, takeovers, and Internet billionaires. At the same time, the Internet search engine population diversified. Newcomers appeared that applied more advanced technologies, such as HotBot in 1995. Also, there was the emergence of "meta search" engines, such as Dogpile, "natural language" search engines, such as AskJeeves, and multimedia searchers. One type of new entrant seemed to particularly proliferate, namely search engines offering their service in a specific language or targeting a particular geographical area. In fact, from 1997 onward most new search engines specialized in finding pages from a specific domain or language area.

The question why most search engine entrepreneurs chose a specialist strategy is intriguing. Especially on the Internet, there seems to be much to gain by aiming at a broad audience, rather than by specialization. Before search services can be profitable, massive numbers of hits have to be generated, and this seems to be easier for services aiming at the general market than for specialized services.

New entrepreneurs, however, may have been hesitant to compete with the large and established players that traditionally serve the general market. Economies of scale are often decisive in competition, and in the search engine market scale economies are substantial (Shapiro and Varian 1999).

But then again, by adopting a specialist strategy, many search engines have willfully surrendered to being small; and smallness is a property that is a hazard to survival in an environment where scale economies dominate.

---

[6] The fact that the first search engine home pages were located in the personal directories of people working at computer science departments, illustrates this observation.

[7] Source: `http://www.searchenginewatch.com`.

The Internet search engine market is not the only market where an—unlikely—emergence of small specialists has been observed (Swaminathan 1995; Swaminathan and Carroll 1995). According to organizational ecologists, such observations show that environments in which scale economies dominate are not hostile, but instead advantageous to small specialists. The following section will summarize the general claims of organizational ecology, as well as some of its more specific claims that apply to the population of Internet search engines.

5.3 Organizational ecology

Organizational ecology investigates the dynamics of organizational populations, or industries,[8] from a Darwinian point of view (Hannan and Freeman 1977; Hannan and Freeman 1989). Organizations are seen as relatively inert with respect to environmental changes. Environmental selection is seen as the main determinant of the diversity of organizational forms, while adaptive behavior is not (Hannan and Freeman 1984).

The fact that adaption to (changing) environments has such a secondary role in organizational ecology puts the theory at odds with classical economic theory, where rational behavior of the *homo economicus* is an indispensable assumption.

Both the refreshing view on economical processes, and the fact that a huge body of empirical evidence is collected to sustain its claims, has made organizational ecology a widely recognized theory among organizational sociologists (Knoke 2000), and has even gained respect among economists (Jovanovic 2001) who have traditionally claimed organizational markets as a scientific domain.

Within organizational ecology, two later developed theories are considered to be the most important parts: density dependence and resource partitioning (Carroll and Hannan 1995). Density dependence explains the growth and decline of organizational populations as determined by competition and legitimation (Hannan and Carroll 1992). Resource partitioning explains differentiation in organizational populations, and relates differentiation to changing size distributions (Carroll 1985). The two theories are unrelated in the sense that none of them builds on the other part in any of its explanations (Carroll and Hannan 1995).

5.3.1 Resource partitioning theory

Resource partitioning theory is one of a small group of sub-theories of Organizational Ecology that explain "segregation processes" (Carroll and Hannan 2000,

[8] The domain of organizational ecology consists of populations of organizations in modern, western societies. A population of organizations contains organizations of similar forms. Examples of populations studied in organizational ecology are newspapers, car manufacturers, labor unions, and day care centers.

p.264). The segregation process under consideration in resource partitioning theory concerns the emergence of small specialist organizations in an originally generalist population. The theory's main claim is that "as market concentration rises, the founding rates of specialist organizations will rise and the mortality rates of specialists will fall." This claim contradicts a theorem of classical economy that states that market concentration reduces life chances of small organizations, specialist or other, by putting up barriers to entry to organizations of smaller size (Schmalensee and Willig 1989).

Resource partitioning theory relates three phenomena to the specialist proliferation. The first is increasing concentration. Concentration is usually defined as the ratio of the aggregated size of the largest three[9] firms in a population and the aggregated size of all firms (Shepherd 1987, pp.563-564). Increasing concentration is due to a domination of scale economies. The second phenomenon is the outflow of generalist organizations, especially among the smaller generalists. The third phenomenon is the "move" of the remaining generalists to the market center—the area where most consumers are to be found. The third phenomenon causes the resources to become "partitioned" between large generalists that aim for central resources, and small specialists, aiming for peripheral resources. Consequently, competition between large organizations and small organizations will decrease.

Resource partitioning theory claims that the life chances of small specialists organizations will improve, if the three phenomena described above are observed in a market. The theory is not very outspoken about whether, and how, the phenomena themselves interrelate, although increasing concentration is tentatively seen as the cause of the outflow of small generalists (Carroll and Hannan 1995, pp.215-221).

Later versions of resource partitioning mention a fourth phenomenon that may stimulate specialists life chances: an increasing "dimensionality of the resource space." According to resource partitioning theory, the resource space is defined by taste dimensions. If tastes become more sophisticated, the number of dimensions of the resource space increases, which is hypothesized to open up empty niche pockets where only specialists fit into (see also Péli and Nooteboom (1999)).

The phenomena described in resource partitioning have been witnessed in, among others, the international automotive industry, the Italian banking industry (Freeman and Lomi 1994), the Dutch newspaper industry (Boone, van Witteloostuijn, and Carroll 2002), and the United States beer brewery industry (Carroll and Swaminathan 2000).

In Chapter 3 the structure of the theory of resource partitioning was studied by the use of formal logic. The aim of this study was to disambiguate resource partitioning theory, to determine the precise relation between the phenomena

[9]Some measures take the ratio of the sizes of the largest four to eight firms, to the size of the entire industry.

described, and to produce a sound and consistent reading of the theory that explained specialist proliferation in a sensible way. The result was a translation of the theory in first-order logic.

The formalized theory says that (1) the number of generalist organizations in the market decreases if scale economies dominate, and that (2) consequently, the remaining generalist focus on the center of the market. This (3) lowers the competitive pressure on specialist organizations in the periphery, which, (4) in turn, improves their life chances. Although concentration is a likely consequence of scale economies, it has no explanatory role in the argument.

Increasing dimensionality of the resource space is, according to the formalization, unrelated to the other explanatory phenomena, but given a number of stringent conditions about the structure of the resource base, it provides an explanation for improving life chances for specialists in its own right.

By predicting a positive relation between scale economies and specialist life chances, resource partitioning provides a possible explanation for the differentiation in the population of search engines. However, resource partitioning theory predicts the positive relation between scale economies and specialist proliferation only for stable environments, that is, for environments where the carrying capacity, *i.e.*, the capacity to sustain organizations, has been reached, and where competition neither de- or increases (see, for example, **BA 1** in Chapter 3 of this thesis).

In the case of the search engine industry, the proliferation of specialists occurred in a rapidly growing environment. To such environments, Carroll and Hannan's theory of resource partitioning does not apply. In order to find out whether resource partitioning processes have nevertheless induced the observed specialist proliferation, we try to generalize resource partitioning theory, so that it can be applied to growing industries, as well.

As a starting point for this exercise, we take density dependence theory. This theory explains growth trajectories of organizational populations. Hannan and Carroll also alluded to linking density dependence to resource partitioning theory (Hannan and Carroll 1992, p.48), but an an actual attempt to link both theories has never been made.

5.3.2 Density dependence theory

Density dependence theory claims that the life chances of organizations in a population are dependent on the population size, *i.e.*, the number of organizations, in that population. Increasing population size increases both the *legitimation* of the pertaining organizational form and the overall *competition* within the population. In density dependence theory, and in organizational ecology in general, the term legitimation is used to denote the so-called "taken-for-grantedness" of an organizational form (Carroll and Hannan 2000, p.223). Legitimation has a negative effect on organizational disbanding rates and a positive effect on founding rates, whereas competition has a positive effect on disbanding rates and a negative effect on founding rates. Founding and disbanding rates together determine the growth of the population. Mergers, transitions, and takeovers are not being taken into account in the current theory.

The main propositions of density dependence have mathematical counterparts. These are qualitative in the sense that the main properties of the functional forms are specified, but the functional forms themselves are not defined. Density dependent population dynamics have been witnessed in a number of industries, among which the semiconductor industry (Hannan and Freeman 1989), the American newspaper industry (Carroll and Hannan 1989b), and the American brewing industry (Carroll and Swaminathan 1991).

Density dependence theory does not distinguish different types of organizations. Life chances are assumed not to vary over individual organizations, whether they are large, small, generalist, specialist, old or young. This is reflected in density dependence's measure of competition, which only considers the size of the population. Density dependence's measure of competition seems incompatible with resource partitioning theory, which considers many different kinds of organizations.

Chapter 4 presented a measure of competition which connects competition at the micro level of individual relations to competition at the macro level of industries. By linking competition at different levels, the measure may also link theories about competition at different levels, such as density dependence and resource partitioning theory. For convenience, in the following section we will shortly present the measure. For a more detailed account of the measure, we refer to Chapter 4.

5.3.3 Competition

Let us start out with explicating a number of assumptions that we make about the domain to which the measure applies. Note that the current section only summarizes information from Chapter 4, and does not introduce new notions.

We assume each organization in a market to have a niche. A niche is a set of resources.[10] Resources are all objects that organizations need for survival, such as consumers, materials, goodwill, and money. We assume each resource to be a social actor, or to be controlled by a social actor. Organizations try to obtain resources by making appeals to (controllers of) resources. If an organization makes an appeal to a resource, the resource is in the niche of the organization. The number of resources in the niche is the *niche size*. We assume all resources to have similar values.

If two firms make an appeal to the same resource, the firms entertain a com-

[10] The definition of niche as a set of resources is a simplification compared to the notion of niche that is employed by organizational ecology. See the discussion of Chapter 4.

petitive relation. The intensity of this relation is determined by the number of resources that the firms make a concurrent appeal to.

To re-iterate the measures of competition, let v_i and v_j be the size of the niches of organizations i and j, and let v_{ij} be the number of resources that i and j make a concurrent appeal to. Let N be the number of firms in the industry. Then:

- v_{ij} is the competitive intensity between organizations i and j
- $c_{ij} = \frac{v_{ij}}{v_i}$ is the competitive pressure of organization j on i
- $c_i = \sum_{j=1, j \neq i}^{N} c_{ij}$ is the competitive pressure on organization i by all competitors.
- $C = \sum_{i=1}^{N} c_i$ is the industry competition.

Measure v_{ij} is also called *niche intersection*. Measure c_{ij} is similar to what Podolny et al. (1996) call "(asymmetric) niche overlap," whereas c_i is similar to what they call "niche crowding."

C can be rewritten as:

$$C = (N^2 - N) \cdot \overline{c_{ij}} \tag{5.1}$$

where $\overline{c_{ij}}$ denotes the mean value of all c_{ij}.

Equation 5.1 implies that, under *ceteris paribus* conditions, industry competition increases with the number of firms in the industry, N, at an increasing rate. This relation between N and industry competition is exactly as it is defined by density dependence theory (Hannan and Carroll 1992, p.40). Therefore, we assume that we can apply our measure of competition to density dependence theory. Besides industry competition, a key determinant of founding and disbanding rates in density dependence theory is legitimation.

5.3.4 Legitimation

Legitimation[11] denotes the degree in which an organizational form is "taken for granted" by relevant actors (*i.e.* controllers of resources). In organization theory legitimation is put forward as a crucial condition for founding; a legitimate form can be readily visualized by entrepreneurs (Hannan and Carroll 1992, p.36); it enables them to found organizations, hence to benefit from existing infrastructure, and it facilitates access to resources from the environment (Hannan and Freeman 1989, p.69). Legitimation also lowers the disbanding of organizations.

[11] In accordance with organizational ecology, we use the noun "legitimation," *i.e.*, the act of legitimizing, instead of "legitimacy," the result of legitimation.

Organizations have to “boost” legitimation of their form themselves.[12] When the first organization of a new form appears, some relevant actors in its environment might judge its form legitimate. More organizations on the scene increase the number of actors that take the form for granted, which, in turn, increases the total legitimation of the form.

Organizational ecology assumes later entrants in the market to have less impact on the legitimation of the form. Therefore, legitimation, L, is assumed to increases with population size, N, at a decreasing rate:

$$L = l \cdot N^{-\alpha} \tag{5.2}$$

where $l > 0$ and $\alpha > 0$.

Estimating legitimation

We need to keep in mind, that the relation between legitimation and population size as described in Equation 5.2 is not more than an assumption based on empirical generalizations. In specific cases, legitimation of a form may rise or fall without changing population size, *e.g.*, due to changing legislation, conjuncture, or fashionability. In the end, legitimation of a form is determined only by the number of relevant actors that take the form for granted.

To be able to determine the degree of legitimation with more precision than Equation 5.2 does, we want to estimate the number of relevant legitimizing actors.

Let us assume that an actor is relevant for an industry if he controls resources that are in the resource base, R, of an industry. Furthermore, let us assume that actors who take the form for granted are appealed to by one or more of the firms in the industry. That is, relevant actors are—or are in control of—resources in the fundamental niche of one or more of the firms. Then, legitimation for an organizational form is proportional to the total number of resources that is in the niche of one or more firms.

Simply adding all niche sizes would lead to an overestimation of the total number of resources; resources in the intersection of two firms would be counted twice. In contrast, adding all niche sizes and subsequently subtracting all intersections, would underestimate legitimation.[13]

We can estimate the total number of resources that the industry appeals to if we know the size of the resource base, and the portion of the resource base that is *not* covered by the niches of any of the organizations. The portion of the resource base not covered by the niche of firm i can be expressed as:

$$1 - \frac{v_i}{R}$$

[12] A new organization, especially if it is the first of a kind, needs some resources of its own to survive an initial period before it has established ties with suppliers and clients. Low mortality just after entry reflects the starting capital of firms (Brüderl, Preisendörfer, and Ziegler 1992).

[13] If a resource is in n niches, it is in $\frac{n^2-n}{2}$ niche intersections. So, the resource will be counted $n - \frac{n^2-n}{2} = \frac{3n-n^2}{2}$ times. For $n > 2$, the resource will be counted 0 times or less.

For two random firms, the portion of the resource base not covered by both their niches can be estimated as the product of the portions that both niches do not cover. So, for organizations i and j,

$$(1 - \frac{v_i}{R})(1 - \frac{v_j}{R})$$

The portion of the resource base not covered by the niches of any of the organizations is the product of all the individual portions. The complement of this product (that is, 1 minus the product) is the portion of the resource base that *is* covered by the niches of all organizations. So, the number of resources appealed to by the population, *i.e.*, legitimation, can be estimated as:

$$L \propto R \cdot (1 - \prod_{i=1}^{N}(1 - \frac{v_i}{R})) \tag{5.3}$$

remember that $\propto$ indicates proportionality. We will apply the legitimation measure expressed in Equation 5.3 instead of organizational ecology's original measure for legitimation. Notice that Equation 5.3 implies that legitimation is maximal if all resources are included in at least one of the organizations' niches.

Figure 5.1 shows the canonical growth trajectory of legitimation with population size according to Equation 5.3.[14] The figure shows that legitimation increases

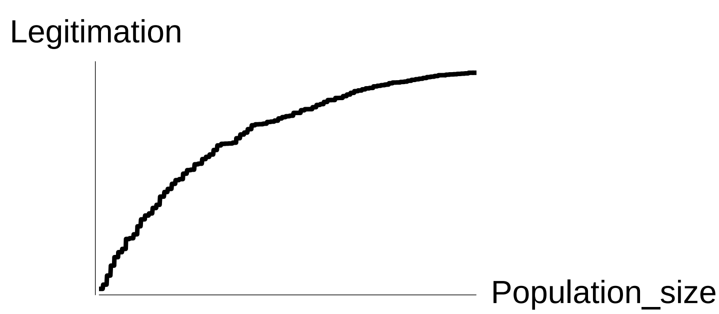


Figure 5.1: Legitimation: portion of resource base coverage

with population size at a decreasing rate. This should be understood as follows: in small industry populations, firms that enter the population are likely to appeal to resources that no other firm appeals to, and thus, to increase legitimation for the organizational form. When a population has grown large, most resources are

[14]Figure 5.1 is based on the following parameters: $R = 100$, N goes from 1 to 100, and v_i is randomly distributed between 0 and 6.

already appealed to, and new firms are less likely to increase legitimation. The relation between legitimation and population size as specified in Equation 5.3, is conform the relation defined in density dependence theory (Hannan and Carroll 1992, p.42).

We assume legitimation to be equal for all firms in the industry. In that respect, legitimation differs from competitive pressure, that varies over the organizations in the industry. The reason for this difference is that legitimation pertains to the *form* of the organizations. An industry, by definition, consists of organizations of the same form; hence, with the same legitimation.

5.3.5 Density dependence at work

Density dependence theory relates industry competition and legitimation to founding and disbanding rates. Industry competition is claimed to have a negative relation to founding rates; it reduces the number of organizations that is founded. Conversely, legitimation increases founding rates. Also, industry competition increases mortality, whereas legitimation reduces mortality.

The relations between industry competition and legitimation on the one hand, and founding rates, F, and mortality rates, M, on the other, are taken to be proportional. The definition is as follows:

$$F \propto \frac{L}{C}, \quad M \propto \frac{C}{L} \tag{5.4}$$

Density dependence predicts a non–monotonous, ∩-shaped relation between population size and founding rate and a non–monotonous, ∪-shaped relation between population size and mortality rate.

5.4 Connecting resource partitioning and density dependence theory

Now that we have discussed both resource partitioning and density dependence theory, we aim to connect both theories, generating a theory that can explain resource partitioning processes in changing competitive environments.

Let N_t be the industry size at time t. Let F_t be the founding rate at t, that is, the number of foundings in the time interval between $t-1$ and t. Let M_t be the mortality rate at t, and let L_t and C_t be the legitimation and the industry competition at t. Density dependence theory assumes that competition increases with population size at an increasing rate. Therefore, competition will also increase with t at an increasing rate. Similarly, L_t will increase at an decreasing rate. Density dependence theory furthermore assumes founding rates to be positively related to legitimation and negatively related to competition. Mortality rates are positively related to competition and negatively related to legitimation.

Finally, as a consequence on founding and mortality rates, population size follows an S-shaped trajectory.

Assumptions 1 to 4 and Claim 1 summarize the argument of density dependence theory. Let f'_t denote the first derivative of function f to t.[15,16]

1. Assumption. *$C'_t > 0$ and $C''_t > 0$ (C_t increases at an increasing rate)*

2. Assumption. *$L'_t > 0$ and $L''_t < 0$ (L_t increases at a decreasing rate)*

3. Assumption. $F_t \propto \frac{L_t}{C_t}$

4. Assumption. $M_\tau \propto \frac{C_\tau}{L_\tau}$

1. Claim. *Before some point of inflection $t = \lambda$, it holds that $N'_t > 0$ and $N''_t > 0$. After $t = \lambda$, it holds that $N'_t > 0$ and $N''_t < 0$.*

Resource partitioning theory considers two subpopulations within a population: generalists and specialists. Let N_t^G be the number of generalist organizations at t; N_t^S is the number of specialists. We assume organizations two be either generalist or specialist, so:

1. Background assumption. $N_t = N_t^G + N_t^S$

Let F_t^G and M_t^G be founding and mortality rates of the generalist subpopulation, and F_t^S and M_t^S be founding and mortality rates of the specialist subpopulation.

C_t^G is the total competitive pressure on the generalist subpopulation, whereas C_t^S is the total competitive pressure on the specialist subpopulation. So:

2. Background assumption. $C_t = C_t^G + C_t^S$

The competitive pressure on the generalist subpopulation is induced by both generalists and specialists. Let C_t^{GG} be the total pressure on generalists by generalists, and C_t^{GS} the total pressure on generalists by specialists. So:

3. Background assumption. $C_t^G = C_t^{GG} + C_t^{GS}$

Similarly, C_t^{SS} and C_t^{SG} are the total competitive pressures on specialists by specialists and generalists respectively, and:

[15] Density dependence's original assumptions take derivations to N instead of to t. By taking derivations to t we simplify the model. This simplification does not alter the original assumptions up to the moment where the carrying capacity of the population is reached, as before that moment, N monotonically increases with time. For a discussion of these and related issues, see (Péli 1993).

[16] Notice that founding rates are defined on t, wheras mortality rates are defined on τ, to indicate that both functions may run accross different time axes.

4. Background assumption. $C_t^S = C_t^{SS} + C_t^{SG}$

Let $\overline{c}_t^{gG}$ be the mean competitive pressure on generalists by all other generalists, so $\overline{c}_t^{gG} = \frac{C_t^{GG}}{N_t^G}$. Similarly, we define:

- $\overline{c}_t^{gS} = \frac{C_t^{GS}}{N_t^G}$
- $\overline{c}_t^{sS} = \frac{C_t^{SS}}{N_t^S}$
- $\overline{c}_t^{sG} = \frac{C_t^{SG}}{N_t^S}$

Let $\overline{c}_t^{gg}$ be the mean competitive pressure imposed on generalists by means of a single generalist/generalist competitive relation. So, $\overline{c}_t^{gg} = \frac{C_t^{gG}}{N_t^G - 1}$ Similarly, we define:

- $\overline{c}_t^{gs} = \frac{C_t^{gS}}{N_t^S}$
- $\overline{c}_t^{ss} = \frac{C_t^{sS}}{N_t^S - 1}$
- $\overline{c}_t^{sg} = \frac{C_t^{sG}}{N_t^G}$

As discussed in Section 5.3, resource partitioning theory relates three phenomena to specialist proliferation: (1) concentration, caused by dominating scale economies, (2) a decreasing number of generalists, and (3) a generalist move to the center. We assume the relations between these phenomena, as well as their relation to specialist proliferation, to be defined by the logical formalization in Chapter 3.

We will ignore for now the increasing dimensionality of the resource base, which has been mentioned as an additional cause for specialist proliferation. In Section 5.7 we will revisit it.

Resource partitioning's main claim is that the life chances of specialist organizations increase if scale economies dominate. However, density dependence theory claims that the life chances of *all* organizations increase until the carrying capacity is reached. So, in industries where the carrying capacity is not reached, density dependence makes resource partitioning's main claim redundant. Therefore, for growing populations, the claim should be restated: due to scale economies, the life chances of specialists will become better than those of generalists.

Resource partitioning theory's first assumption is that the number of generalists decreases if scale economies dominate. If generalist numbers decrease in an otherwise growing population, the number of specialists automatically increase—there are no other organizations than generalists and specialists in the population. An increase of specialists, in turn, implies that life chances of specialists are better than those of generalists, which immediately implies the conclusion of the main claim of the theory. So, for growing populations, the assumption that the number of generalists decreases is too strong. A weaker, and for growing populations more appropriate, assumption is that the *growth rate* of the generalist subpopulation decreases:

5. ASSUMPTION. *If scale economies dominate, then* $N_t''^G < 0$

The next assumption of resource partitioning says that generalist organizations move to the center of the resource base if the founding rate of the generalist subpopulation decreases. So, generalist niches will, increasingly, be located in one particular area of the resource base. If generalist niches converge, the mean competitive pressure imposed by generalists on their generalist peers will increase.

6. ASSUMPTION. *If* $N_t''^G < 0$*, then* $\overline{c}_t'^{gg} > 0$

If generalists move to the center of the resource base, they move out of the periphery. By doing so, the competition between generalists and specialists decreases—more precisely, the mean competitive pressure of generalist firms on specialist firms decreases.

7. ASSUMPTION. *If* $\overline{c}_t'^{gg} > 0$*, then* $\overline{c}_t'^{sg} < 0$

Subsequently, because the competitive pressure of generalists on specialists decreases, the mean competitive pressure on specialists decreases.

8. ASSUMPTION. *If* $\overline{c}_t'^{sg} < 0$*, then* $\overline{c}_t'^{S} < 0$

If the mean competitive pressure on specialists decreases, specialists life chances will improve, so, the growth rate of the specialist subpopulation increases.

9. ASSUMPTION. *If* $\overline{c}_t'^{S} < 0$*, then* $N_t''^S > 0$

From the assumptions of both density dependence and resource partitioning theory, the following theorems[17] can be derived[18]:

5.4.1. THEOREM. *If scale economies dominate, then* $\overline{c}_t'^{gg} > 0$ *and* $\overline{c}_t'^{sg} < 0$

5.4.2. THEOREM. *If scale economies dominate, then* $N_t''^G < 0$ *and* $N_t''^S > 0$

Note that the assumptions of resource partitioning are all conditional. That means that they are falsified by an empirical domain, if the condition (the antecedent) is true in the domain, but the implication (the consequent) is false. For example, if, in a market, scale economies dominate, but the growth rate of the generalist subpopulation does not decrease, then Assumption 5 does not hold. If all Assumptions are true for some t, then the theorems are necessarily true for t.

Before we test the assumptions in Section 5.6, in the following section we will present the data and the empirical domain: the market of Internet search engines.

[17]Note that we call the main claim of density dependence a "claim," whereas we call the main claims of resource partitioning "theorems." The difference is that the theorems are proven—from Assumptions 5 to 9—whereas the claim originates in (Hannan and Carroll 1992), and is, as far as we know, not proven.

[18]See Appendix I for a proof.

5.5 Data

This study uses data on the international Internet search engine market. It considers the full history of the search engine market, starting with its inception in 1993, and ending in October 2000. Data has been collected for each quarter year, amounting to 28 assessments.

5.5.1 Source of the data

The data has been provided by Internet search engine Northernlight.[19] The data set consists of the results, as returned by Northernlight, on a series of approximately 265,000 automatically generated queries, concerning the prevalence of unambiguous references to Internet search engines, in publicly accessible, dated, Internet documents.

Northernlight was chosen as a source for data collection because of four reasons. First, Northernlight has a large index that covers a relatively large part of the Internet (Lawrence and Giles 1999). The coverage of Northernlight is among the best of the search engines (Hawking and Craswell 2001). Second, the number of results that Northernlight reported on was stable, in the sense that on a query that was repeated after a short period of time (say, a few hours), the same number of results is returned. Other large search engines show varying numbers of results on the same query. Third, the number of results that Northernlight returns is both precise and reliable. Some of the other large search engines provide only an approximation of the number of results. Northernlight, in contrast, gives a precise number, and moreover, actually displays all results, such that precision and reliability can be checked. Most search engines display a maximum of 200, or 1000 results.

The fourth reason to use Northernlight as a source of data rather than other large search engines, was because it is able to process complex queries correctly.[20] Other large search engines do not allow for complex queries, or generate inconsistent results.

5.5.2 Identifying the industry members

The first step in our empirical study was to identify the industry of search engines, both past and present. This was done by performing an exploratory study of an number of Internet sources, such as the Search Engine Watch,[21] the Internet

[19] `http://www.northernlight.com`.

[20] Examples of "complex queries" are queries that employ boolean operators, such as AND, OR, and NOT, and that ask for results restricted to specified time intervals.

[21] Address: `http://www.searchenginewatch.com`.

Archive,[22] and the Search Engine Guide.[23] The result of this study was a list of 137 Internet search engines that were mentioned in one of these sources. Most of these search engines were still active, but some had already been shut down. When possible, we have established the origin, search domain, target domain, language of communication of, and additional services provided by the engines. We did this by visiting the engines, and by actually using their search service.

Our aim was to establish a complete list of all search engines existing somewhere between 1993 and 2000. It is possible, however, that some search engines have escaped our notice. Some search engines may have come into existence and withered away without ever getting noticed by one of our sources, and consequently, by us. Nevertheless, we believe that the group of search engines that constitute our empirical population, comprises a significant part, and thus, a representative sample, of the actual search engine population. The group of search engines that we investigate in this study is one order of magnitude larger than groups investigated in other recent studies (Gandal 2001; Lawrence and Giles 1999). Contrary to the search engine populations from these studies, our population includes many search engines that originate in countries other than the United States, or that offer their services in languages other than English.

5.5.3 Data collection

Using the list of 137 Internet search engines, we automatically generated two types of queries. First, we asked Northernlight for the number of Internet pages that mentioned the URL ("Uniform Resource Locator," see also Section 4.3.3) of one of the search engines from the list. We did this for each quarter between October 1993 and October 2000, so, 28 times.

Second, we asked for the number of Internet pages that mentioned the URLs of each possible pair of search engines from our list of 137. Again, we did this for each of the 28 pertaining quarters.

We collected all our data in 2001, which may imply older pages to be relatively underrepresented in our database. To address this problem, we asked Northernlight for the total number of pages that it has in its index for each of the 28 quarters.[24]

In order to determine whether older pages are proportionally represented in Northernlight's index, we obtained data about the estimated size and growth of

[22] Address: `http://www.archive.org`.

[23] Address: `http://www.searchengineguide.com`.

[24] The expectation that older pages are underrepresented in the index of a search engine such as Northernlight may seem obvious, but in fact is not. Older pages may very well be underrepresented on the Internet itself, because they have a larger chance of being removed, or updated. Search engines, however, do not visit all pages on the Internet on a regular basis, and do not notice most changes. As far as they are concerned, Internet pages remain as they were last encountered. Additionally, older pages have a better chance of being indexed by search engines, just because they have been given a longer period to be discovered.

the internet from another source: "Hobbes' Internet Timeline."[25]

Figure 5.2 illustrates the number of measurements that we have done. We "fed" Northernlight 260,848 queries[26] asking for Internet pages mentioning the URLs of a pair of search engines. Figure 5.2 shows all instances where such a query returned at least one result. Two axes of the figure represent search engine pairs, the third axis represents the yearly quarters.

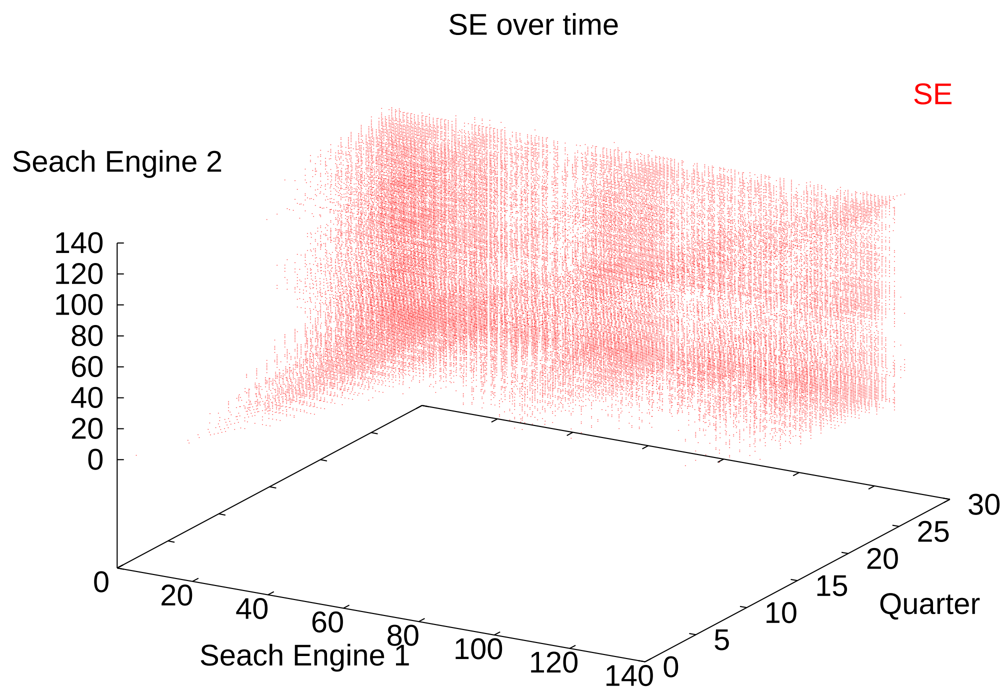


Figure 5.2: Measurements per search engine pair, per yearly quarter

Finally, we visited all 137 search engines to obtain information about their origin, search domain, target domain, language of communication, and complementary services provided on their web site, such as free e-mail, disk space, yellow guide, directory, translations, portal, etc.

[25] Address: `http://www.zakon.org/robert/internet/timeline/`.

[26] The queries were automatically generated and fed to Northernlight over a 10 day period. To not overload Northernlight's search service, we restricted our data collection to non–office hours.

5.5.4 Measurements

In order to test our model, we want to measure competitive relations between—different types of—firms. Thus, we need to establish for each firm in the population whether it is a generalist or a specialist; we need to establish its founding date and its disbanding date; and we need to establish the competitive pressure imposed on it. In order to establish competitive pressure on a focal organization, we need to know the size of its niche, and the degree in which other organizations target its niche.

Niche size

If an author of an Internet page mentions a string (a row of letters) that unambiguously refers to a particular search engine (*e.g.*, "AltaVista.com") the author is likely to be familiar with that particular search engine. To determine search engine niche size, we aim at measuring not the people actually using search engines, but the people familiar with them; in that way we measure fundamental (potential) niches, and not realized niches. Organization ecology stresses the difference between the two: realized niches are considered the result of competition, whereas fundamental niches are the source of competition (Carroll and Hannan 2000).

We measure the size of the niche of search engine i at quarter t, v_{it}, as the number of Internet pages referring to search engine i at t, according to Northernlight; we call this measure $\mathbf{v}_{it}$. We determined $\mathbf{v}_{it}$ for each of the 137 search engines, and for each of the 28 quarters; a total of 3,836 assessments.

Niche intersection

The intensity of competition between two search engines i and j at time t is determined by the intersection of their niches at t, v_{ijt}. We measure niche intersection by the number of Internet pages from quarter t that refer to both search engine i and to search engine j at t.

This number is denoted as $\mathbf{v}_{ijt}$. We measured $\mathbf{v}_{ijt}$ for each possible pair of search engines, for each quarter; a total of 260,848 assessments.

Generalist or specialist

We established five "observational" scores per search engine, (1) country of origin, (2) search domain, (3) target domain, (4) language of communication, and (5) other services. From these scores we determined whether a search engine should be considered generalist or specialist. We regard a search engine generalist if (a) the principle language of operation is English, and (b) either the search domain, or the target domain, is not restricted. Otherwise a search engine is regarded specialist.[27]

The fact that search engine i is a generalist is denoted as $\mathbf{gen}_i = 1$. The fact that j is a specialist is denoted as $\mathbf{spc}_j = 1$. All firms are either generalist or specialist. We apply the organizational ecological assumption that organizations cannot change their organizational form radically enough to transform from generalist into specialist or vice versa (Freeman and Hannan 1983; Hannan and Freeman 1984).

Population size and vital rates

To measure population size N_t, we have to establish the number of search engines that is "in business" at t. Let $\mathbf{x}_{it}$ denote that search engine i exists at t.

We consider a search engine founded if its niche size exceeds 10. We apply this threshold to limit the possibility that our measurements are based on a small number of falsely dated Internet pages. Let $\mathbf{f}_{it}$ denote that search engine i is founded at t. We say $\mathbf{f}_{it} = 1$ if $\mathbf{v}_{it} > 10$ and $\mathbf{x}_{it-1} = 0$, and $\mathbf{f}_{it} = 0$ otherwise.

We consider a search engine disbanded if its niche size becomes smaller then 10, for at least two quarters in a row. Let $\mathbf{m}_{it}$ denote that search engine i has disbanded at t. We say $\mathbf{m}_{it} = 1$ if $\mathbf{v}_{it} < 10$, and $\mathbf{v}_{it-1} < 10$, and $\mathbf{x}_{it-1} = 1$. Otherwise, $\mathbf{m}_{it} = 0$.

A search engine exists if it is founded, and has not yet been disbanded. We say $\mathbf{x}_{it} = 1$, if $\mathbf{f}_{it} = 1$ or if both $\mathbf{x}_{it-1} = 1$ and $\mathbf{m}_{it} = 0$. Notice, that the definition of $\mathbf{x}_{it}$ is recursive. To avoid circularity, we must add the condition that for all i, $\mathbf{x}_{it_0} = 0$, that is, at t_0 no search engine is in business.

Using $\mathbf{x}$, $\mathbf{f}$, and $\mathbf{m}$, we define measures for population size, and for founding rates and mortality rates. That is:

$$\mathbf{N}_t = \sum_{i=1}^{137} \mathbf{x}_{it}$$

The founding rate is the number of firms founded at time t. Similarly, mortality rate is the number of firms disbanding at t:

$$\mathbf{F}_t = \sum_{i=1}^{137} \mathbf{f}_{it} \qquad \mathbf{M}_t = \sum_{i=1}^{137} \mathbf{m}_{it}$$

Subpopulation size and vital rates

The size of the generalist subpopulation is simply the number of search engines in the industry that are generalist. Similarly, the size of the specialist subpopulation

[27] As an example, search engine WebTop, which specializes in finding European pages, but targets a general public in English, is regarded by us to be a generalist. Conversely, French search engine Voila.fr, which specializes in finding European pages, and is operated in French, is regarded by us to be a specialist.

is the number of search engines in the industry that are specialist:

$$\mathbf{N}_t^G = \sum_{i=1}^{137} \mathbf{gen}_i \cdot \mathbf{x}_{it} \qquad \mathbf{N}_t^S = \sum_{i=1}^{137} \mathbf{spc}_i \cdot \mathbf{x}_{it}$$

Generalist and specialist founding rates are:

$$\mathbf{F}_t^G = \sum_{i=1}^{137} \mathbf{f}_{it} \cdot \mathbf{gen}_i \qquad \mathbf{F}_t^S = \sum_{i=1}^{137} \mathbf{f}_{it} \cdot \mathbf{spc}_i$$

Generalist and specialist mortality rates are:

$$\mathbf{M}_t^G = \sum_{i=1}^{137} \mathbf{m}_{it} \cdot \mathbf{gen}_i \qquad \mathbf{M}_t^S = \sum_{i=1}^{137} \mathbf{m}_{it} \cdot \mathbf{spc}_i$$

Legitimation

We defined legitimation as the estimated coverage of the resource base by the niches of all firms. The estimation was based on the sizes of all niches $\mathbf{v}_{it}$, and the size of the resource base. The size of the resource base at time t we indicate by the number of pages on the Internet that were created in quarter t, as available to Northernlight. We obtained that number by asking Northernlight for all pages created in quarter t that did *not* contain the—meaningless and unlikely—character string *fsdgdhbdnbjgbsdkjgbskggetc*.

Let $\mathbf{R}_t$ denote the size of the resource base at t, then:

$$\mathbf{L}_t = \mathbf{R}_t \cdot (1 - \prod_{i=1}^{137} (1 - \frac{\mathbf{v}_{it} \cdot \mathbf{x}_{it}}{\mathbf{R}_t}))$$

Competition

Industry competition is the sum of the competitive pressure exerted by all firms in the industry on all other firms in the industry. So,

$$\mathbf{C}_t = \sum_{i=1}^{137} \sum_{j=1, j\neq i}^{137} \frac{\mathbf{v}_{ijt}}{\mathbf{v}_{it}} \cdot \mathbf{x}_{it} \cdot \mathbf{x}_{jt}$$

The total competitive pressure exerted on the generalist subpopulation is:

$$\mathbf{C}_t^G = \sum_{i=1}^{137} \sum_{j=1, j\neq i}^{137} \frac{\mathbf{v}_{ijt}}{\mathbf{v}_{it}} \cdot \mathbf{x}_{it} \cdot \mathbf{x}_{jt} \cdot \mathbf{gen}_i$$

Similarly, the total competitive pressure exerted on specialists:

$$\mathbf{C}_t^S = \sum_{i=1}^{137} \sum_{j=1, j\neq i}^{137} \frac{\mathbf{v}_{ijt}}{\mathbf{v}_{it}} \cdot \mathbf{x}_{it} \cdot \mathbf{x}_{jt} \cdot \mathbf{spc}_i$$

The mean competitive pressure exerted on generalists by generalists, is:

$$\overline{\mathbf{c}}_t^{gG} = \frac{1}{\mathbf{N}_t^G} \cdot \sum_{i=1}^{137} \sum_{j=1, j \neq i}^{137} \frac{\mathbf{v}_{ijt}}{\mathbf{v}_{it}} \cdot \mathbf{x}_{it} \cdot \mathbf{x}_{jt} \cdot \mathbf{gen}_i \cdot \mathbf{gen}_j$$

Similarly,

$$\overline{\mathbf{c}}_t^{gS} = \frac{1}{\mathbf{N}_t^G} \cdot \sum_{i=1}^{137} \sum_{j=1, j \neq i}^{137} \frac{\mathbf{v}_{ijt}}{\mathbf{v}_{it}} \cdot \mathbf{x}_{it} \cdot \mathbf{x}_{jt} \cdot \mathbf{gen}_i \cdot \mathbf{spc}_j$$

$$\overline{\mathbf{c}}_t^{sS} = \frac{1}{\mathbf{N}_t^S} \cdot \sum_{i=1}^{137} \sum_{j=1, j \neq i}^{137} \frac{\mathbf{v}_{ijt}}{\mathbf{v}_{it}} \cdot \mathbf{x}_{it} \cdot \mathbf{x}_{jt} \cdot \mathbf{spc}_i \cdot \mathbf{spc}_j$$

$$\overline{\mathbf{c}}_t^{sG} = \frac{1}{\mathbf{N}_t^S} \cdot \sum_{i=1}^{137} \sum_{j=1, j \neq i}^{137} \frac{\mathbf{v}_{ijt}}{\mathbf{v}_{it}} \cdot \mathbf{x}_{it} \cdot \mathbf{x}_{jt} \cdot \mathbf{spc}_i \cdot \mathbf{gen}_j$$

The mean competitive pressure exerted by a single generalist/generalist relation, is:

$$\overline{\mathbf{c}}_t^{gg} = \frac{\overline{\mathbf{c}}_t^{gG}}{\mathbf{N}_t^G - 1}$$

Similarly,

$$\overline{\mathbf{c}}_t^{gs} = \frac{\overline{\mathbf{c}}_t^{gS}}{\mathbf{N}_t^S} \qquad \overline{\mathbf{c}}_t^{ss} = \frac{\overline{\mathbf{c}}_t^{sS}}{\mathbf{N}_t^S - 1} \qquad \overline{\mathbf{c}}_t^{sg} = \frac{\overline{\mathbf{c}}_t^{sG}}{\mathbf{N}_t^G}$$

That concludes the operational definitions of our measurements.

5.6 Results

We measured all variables according to the operational definitions of Section 5.5. The results are displayed in Table 5.1 (densities and vital rates) and Table 5.2 (legitimation and competition).

5.6.1 Population size

Figure 4.10 displays the population size in the Internet search engine market, from 1993 to 2000 (note that this figure is displayed in Chapter 4). On first sight, population size seems to follow an S-shaped trajectory, as claimed by density dependence theory. Figure 5.3 shows the first derivative of the population size, the populational growth rate. In this figure, we have added the optimal quadratic regression line[28] for N_t'. The fit of the optimal quadratic regression is

[28] The optimal quadratic regression is: $\hat{N}_t' = (-1.69 + 0.78t + 0.002 \cdot t^2)$ For this regression it holds that $R^2 = 0.51$ and $T = 3.56$.

t		$\mathbf{N}_t$	$\mathbf{N}_t^G$	$\mathbf{N}_t^S$	$\mathbf{F}_t$	$\mathbf{F}_t^G$	$\mathbf{F}_t^S$	$\mathbf{M}_t$	$\mathbf{M}_t^G$	$\mathbf{M}_t^S$
1	12/93	0	0	0	0	0	0	0	0	0
2	03/94	2	2	0	2	2	0	0	0	0
3	06/94	2	2	0	0	0	0	0	0	0
4	09/94	3	3	0	1	1	0	0	0	0
5	12/94	5	5	0	2	2	0	0	0	0
6	03/95	9	9	0	4	4	0	0	0	0
7	06/95	10	9	1	1	0	1	0	0	0
8	09/95	12	11	1	2	2	0	0	0	0
9	12/95	13	12	1	1	1	0	0	0	0
10	03/96	18	16	2	5	4	1	0	0	0
11	06/96	19	16	3	1	0	1	0	0	0
12	09/96	23	17	6	4	1	3	0	0	0
13	12/96	29	18	11	6	1	5	0	0	0
14	03/97	33	19	14	4	1	3	0	0	0
15	06/97	37	19	18	5	1	4	1	1	0
16	09/97	47	19	28	10	0	10	0	0	0
17	12/97	53	20	33	8	1	7	2	0	2
18	03/98	61	22	39	8	2	6	0	0	0
19	06/98	69	23	46	8	1	7	0	0	0
20	09/98	81	25	56	12	2	10	0	0	0
21	12/98	88	26	62	7	1	6	0	0	0
22	03/99	92	27	65	4	1	3	0	0	0
23	06/99	96	27	69	5	0	5	1	0	1
24	09/99	101	28	73	6	1	5	1	0	1
25	12/99	108	28	80	8	1	7	1	1	0
26	03/00	117	28	89	9	0	9	0	0	0
27	06/00	121	29	92	4	1	3	0	0	0
28	09/00	124	29	95	3	0	3	0	0	0

Table 5.1: Population size and vital rates in the Internet search engine market

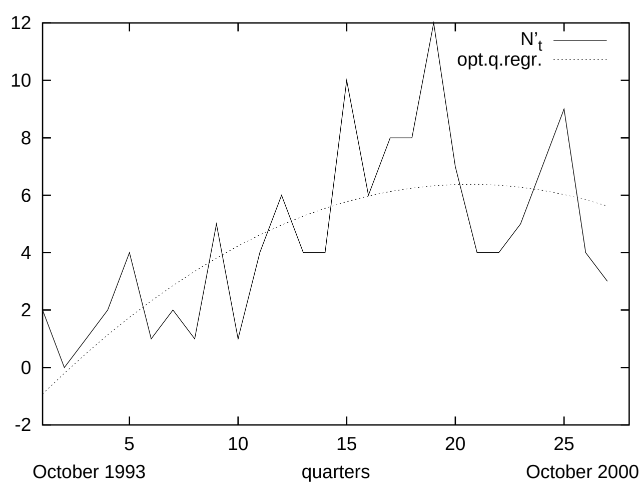


Figure 5.3: Growth rates

t		$\mathbf{L}_t$	$\mathbf{C}_t$	$\mathbf{C}_t^G$	$\mathbf{C}_t^S$	$\overline{\mathbf{c}}_t^g$	$\overline{\mathbf{c}}_t^s$	$\overline{\mathbf{c}}_t^{gG}$	$\overline{\mathbf{c}}_t^{gS}$	$\overline{\mathbf{c}}_t^{sS}$	$\overline{\mathbf{c}}_t^{sG}$	$\overline{\mathbf{c}}_t^{gg}$	$\overline{\mathbf{c}}_t^{gs}$	$\overline{\mathbf{c}}_t^{ss}$	$\overline{\mathbf{c}}_t^{sg}$
1	12/93	0.00	0.00	0.00	0.00	-	-	-	-	-	-	-	-	-	-
2	03/94	0.27	0.00	0.00	0.00	0.00	-	0.00	0.00	-	-	.000	-	-	-
3	06/94	0.26	0.00	0.00	0.00	0.00	-	0.00	0.00	-	-	.000	-	-	-
4	09/94	0.13	0.40	0.40	0.00	0.13	-	0.13	0.00	-	-	.067	-	-	-
5	12/94	0.15	0.60	0.60	0.00	0.12	-	0.12	0.00	-	-	.030	-	-	-
6	03/95	0.61	1.84	1.84	0.00	0.20	-	0.20	0.00	-	-	.026	-	-	-
7	06/95	1.05	5.06	5.06	0.00	0.56	0.00	0.56	0.00	0.00	0.00	.070	.000	-	.000
8	09/95	1.83	10.55	10.55	0.00	0.96	0.00	0.96	0.00	0.00	0.00	.096	.000	-	.000
9	12/95	2.88	12.09	12.09	0.00	1.01	0.00	1.01	0.00	0.00	0.00	.092	.000	-	.000
10	03/96	5.60	20.16	20.16	0.00	1.26	0.00	1.26	0.00	0.00	0.00	.084	.000	.000	.000
11	06/96	9.17	22.16	21.85	0.31	1.37	0.10	1.37	0.00	0.00	0.10	.091	.000	.000	.007
12	09/96	12.89	34.32	31.34	2.97	1.84	0.50	1.84	0.00	0.00	0.50	.115	.001	.000	.029
13	12/96	14.27	37.82	27.95	9.87	1.55	0.90	1.53	0.03	0.05	0.85	.090	.002	.005	.047
14	03/97	19.45	54.92	40.09	14.83	2.11	1.06	2.03	0.08	0.10	0.96	.113	.006	.008	.051
15	06/97	27.01	69.18	42.17	27.02	2.22	1.50	2.10	0.12	0.15	1.35	.117	.007	.009	.071
16	09/97	37.22	96.85	42.30	54.55	2.23	1.95	2.01	0.21	0.38	1.57	.112	.008	.014	.083
17	12/97	66.15	134.01	55.86	78.16	2.79	2.37	2.57	0.22	0.50	1.87	.135	.007	.016	.094
18	03/98	61.12	202.24	75.59	126.65	3.44	3.25	2.92	0.51	1.07	2.18	.139	.013	.028	.099
19	06/98	69.67	227.56	69.07	158.49	3.00	3.45	2.50	0.51	0.99	2.45	.113	.011	.022	.107
20	09/98	95.15	362.67	92.57	270.10	3.70	4.82	3.07	0.64	2.05	2.77	.128	.011	.037	.111
21	12/98	124.66	556.88	136.04	420.84	5.23	6.79	3.76	1.47	2.78	4.01	.150	.024	.046	.154
22	03/99	177.26	663.37	167.87	495.50	6.22	7.62	4.81	1.40	3.12	4.50	.185	.022	.049	.167
23	06/99	246.30	581.30	125.52	455.77	4.65	6.61	3.75	0.90	3.15	3.45	.144	.013	.046	.128
24	09/99	312.22	590.25	131.03	459.22	4.68	6.29	3.86	0.82	2.81	3.48	.143	.011	.039	.124
25	12/99	385.31	645.71	152.11	493.60	5.43	6.17	4.53	0.90	2.54	3.63	.168	.011	.032	.130
26	03/00	528.22	828.59	176.09	652.50	6.29	7.33	5.06	1.22	3.21	4.12	.188	.014	.036	.147
27	06/00	804.73	686.23	164.91	521.32	5.69	5.67	4.60	1.09	2.36	3.31	.164	.012	.026	.114
28	09/00	972.32	789.88	213.15	576.73	7.35	6.07	5.97	1.38	2.39	3.68	.213	.014	.025	.127

Table 5.2: Competition and legitimation in the Internet search engine market

better than that of the optimal linear regression,[29] which suggest that N'_t follows a ∩-shaped trajectory. So, density dependence's main claim is satisfied in the industry of Internet search engines. The question is now, are the canonical growth rate trajectories induced by competition and legitimation, as argued by density dependence theory? In the following two sections we will test whether the explanation provided by density dependence theory applies to our empirical domain. In Section 5.6.2 we will focus on competition and legitimation trajectories, and in Section 5.6.3 we will discuss founding and mortality rates.

5.6.2 Competition and legitimation

Assumption 1 states that industry competition increases at an increasing rate, whereas Assumption 2 states that legitimation increases at a decreasing rate. Figure 5.4 displays both competition and legitimation in the industry of Internet search engines over time. As it turns out, legitimation increases with population

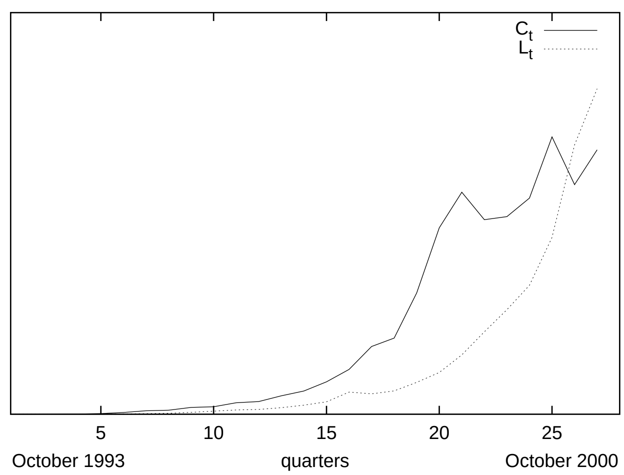


Figure 5.4: Competition and legitimation

size at an increasing rate, rather than at a decreasing rate. Until t_{19}, competition also increases with population size at an increasing rate, but then stagnates, yielding an S-shaped pattern.

So, both Assumptions 1 and Assumption 2 are rejected by the empirical domain. In Sections 5.3.3 and 5.3.4, we showed that *ceteris paribus*, the relations between population size and competition, and population size and legitimation, as proposed by density dependence theory, would be maintained. Apparently,

[29]The optimal linear regression is: $\hat{N}'_t = (1.05 + 0.23 \cdot t)$, with $R^2 = 0.38$ and T=3.98. The F-ratio of the two fits is 5.33, which exceeds the 0.05% significance level at 4.24 (DF, numerator: 1, denominator: 25). The implication is that the quadratic regression is significantly better than the linear regression.

the *ceteris paribus* assumption does not apply to the empirical domain under investigation. Let us take a moment to investigate this conjecture.

Pairwise competition and niche coverage

In the industry of Internet search engines, the increase of competition eventually slows down (Figure 5.4), while population size increases unabated (Figure 4.10). Equation 5.1 implies that in an industry where population size increases and industry competition does *not* increase at an increasing rate, it must be the case that the mean competitive pressure $\bar{c}_{ij}$ decreases. Figure 4.12 (note: this figure is found in Chapter 4) shows the mean competitive pressure over time. We clearly see that after a competitive period in the starting period of the market, $\bar{c}_{ij}$ decreases and stabilizes at a lower level: firms are less competitive in their pairwise relations to fellow industry members.

Figure 5.4 shows that legitimation accelerates. The implication is that at a relatively late stage in the market the search engine form gains notoriety with a critical mass (Schelling 1978, p.91-110) of Internet users. In Figure 5.5a we see that the mean niche size of the search engines increases considerably. In Figure 5.5b, moreover, we see that from t_{17} onward, the growth of the niche sizes even keeps up with the growth of the resource base. If the number of search engines grows, and their niches grow as fast as the resource base, we would expect the resource base to become more "crowded." However, at the same time the mean pairwise niche intersection does not increase. Therefore, it must be the case that the search engines differentiate by covering new grounds and by extending their niches into different directions. This keeps competitive pressure down, and at the same time results in an increasing legitimation. Because of this differentiation process, the *ceteris paribus* conditions of Assumptions 1 and 2 do not hold in the Internet search engine industry, and therefore, the relations described in both assumptions are not found.

Figure 5.6 gives a stylized representation of niche dynamics such as they seem to have occurred in the Internet search engine industry from t_{17} (the last quarter of 1997) onward. The figure shows a growing population with growing niches in a growing resource base. The relative niche intersection stabilizes, amounting into a stable $\bar{c}_{ij}$ and a higher resource base coverage.

5.6.3 Vital rates

Density dependence theory assumes founding rates and mortality rates to be dependent on legitimation and competition. Table 5.3 shows the correlations between two sets of values. On the one hand are the ratio of legitimation and competition, which is assumed to predict founding rates (Assumption 3), the ratio of competition and legitimation, which is assumed to predict mortality rates (Assumption 4), legitimation, and competition. On the other hand are founding

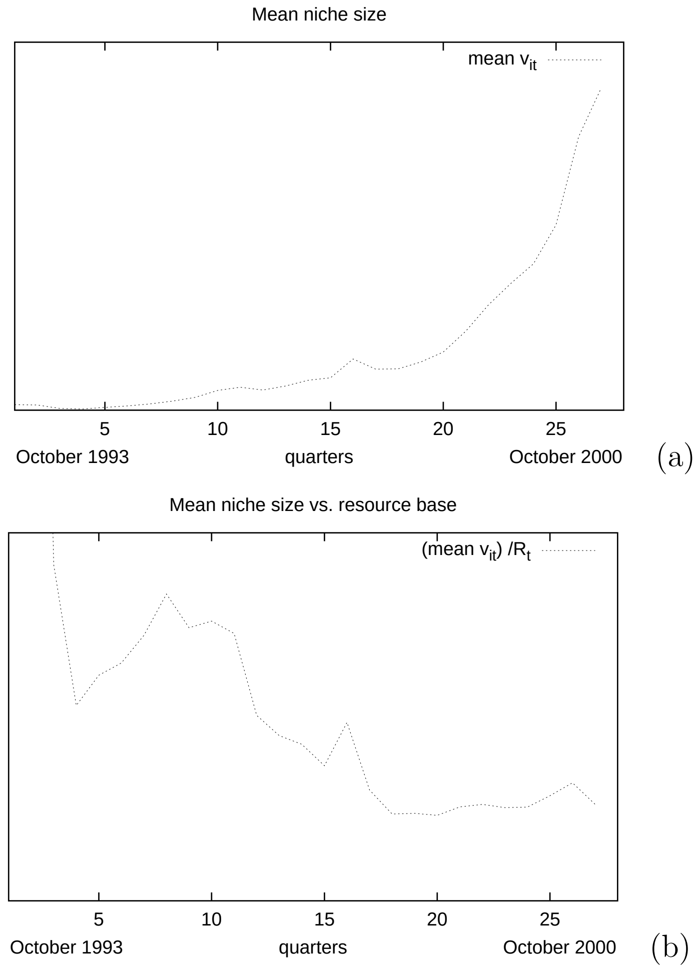


Figure 5.5: Niche size dynamics

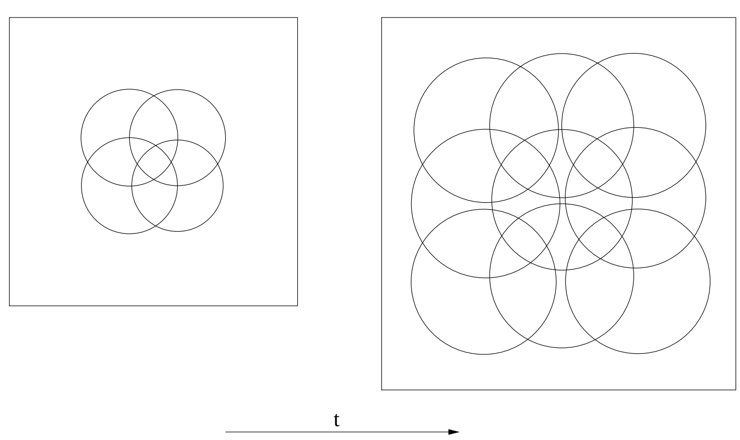


Figure 5.6: Stylized representation of niche dynamics

	$\frac{\mathbf{L}_t}{\mathbf{C}_t}$	$\frac{\mathbf{C}_t}{\mathbf{L}_t}$	$\mathbf{L}_t$	$\mathbf{C}_t$
$\mathbf{F}_t$	0.04		0.20	0.44*
	(0.87)	-	(0.31)	(0.02)
$\mathbf{M}_t$		-0.22	0.09	0.19
	-	(0.27)	(0.64)	(0.32)
$\mathbf{F}_{t+1}$	0.09		0.14	0.31
	(0.68)	-	(0.47)	(0.12)
$\mathbf{M}_{t+1}$		-0.05	0.06	0.21
	-	(0.79)	(0.75)	(0.28)

Table 5.3: Correlations (significance)

and mortality rates at t and at $t+1$. The rates at $t+1$ are added to test for a possible delayed effect of legitimation and competition. It turns out that there is no relation between the founding rates predicted by Assumption 3, and the measured founding rates. The correlation between the predicted and the measured rates is less than 0.1, both for t and $t+1$. Figure 5.7 shows founding rates as predicted by the density dependence model, and the founding rates as measured in the industry of internet search engines.

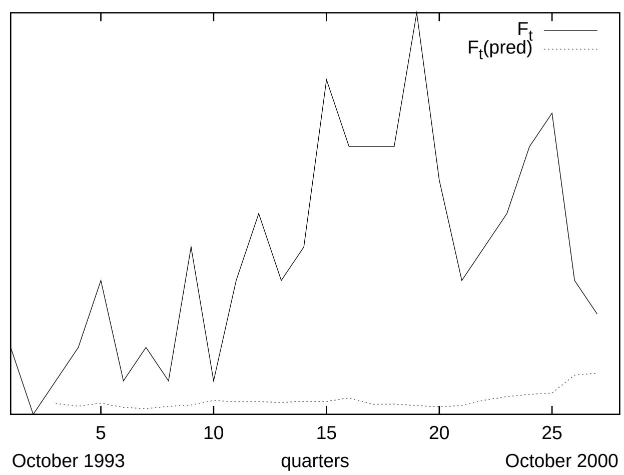


Figure 5.7: Predicted and measured founding rates

With respect to mortality rates, there is also no relation between the predicted and the measured values. The correlation between the predicted and the measured mortality rates turns out to be slightly negative for both t and $t+1$. We should note that these correlations are based on few instances, as we have measured few cases of disbanding firms.

Figure 5.8 shows mortality rates as predicted by the density dependence model, and the mortality rates as measured in the industry of internet search engines.

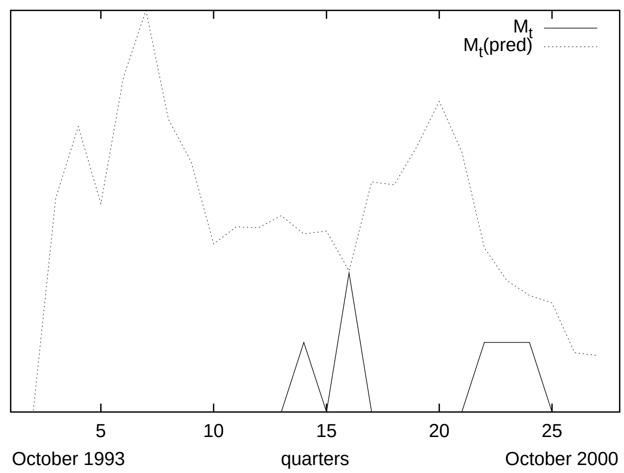


Figure 5.8: Predicted and measured mortality rates

Apparently, in the market of Internet search engines, legitimation and industry competition do not have the predicted relation with vital rates as assumed by density dependence theory. We also considered the correlations between vital rates and competition and legitimation individually (see Table 5.3); we found a positive correlation between competition and founding rates, which, in fact, is the opposite of what was assumed by density dependence theory.

So, Assumptions 3 and 4 of density dependence theory do not apply to our empirical domain. In the previous paragraph we showed that density dependence's Assumptions 1 and 2 do not apply to the search engine industry either. In sum, we found evidence for the main claim of density dependence theory, but we did not find evidence supporting the argument behind the claim.

5.6.4 Generalists and specialists

Figure 5.9 shows the growth trajectories of the generalist and specialist subpopulations in the market of Internet search engines.

Generalists are the first to appear in the market. After a while specialists enter, and soon the specialists constitute a large part of the population. So, the implication of the main claim of resource partitioning theory, specialists vital rates improve relative to those of generalists, is true in our empirical domain. In Figure 5.10, which displays the founding rates of generalist (5.10a) and specialist (5.10b) firms, the relative improvement of specialist founding rates is clearly visible.

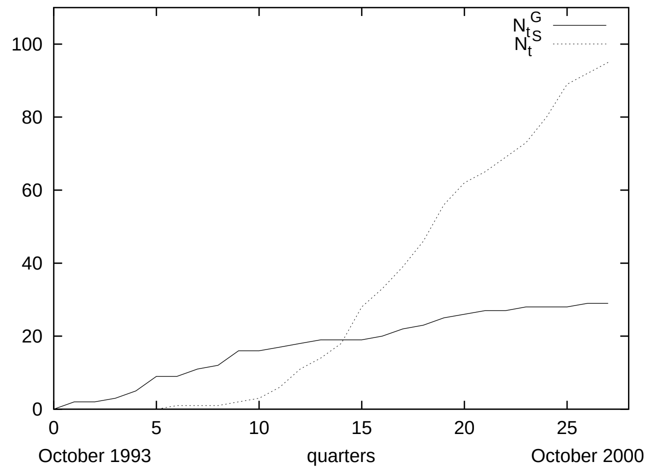


Figure 5.9: Growth of generalist and specialist subpopulations

Although the main implication of resource partitioning theory applies to our empirical domain, we still need to determine whether resource partitioning's explanatory argument applies as well. That is, of all assumptions of resource partitioning, we need to determine whether they hold in our empirical domain.

Generalist founding

Resource partitioning theory starts out by claiming that generalist founding rates decrease if scale economies dominate in a market (Assumption 5). As we argued in the introduction of this chapter, we assume scale economies to dominate in the Internet search engine market. Figure 5.10 shows that, indeed, generalist founding rates decrease in the focal period. A rather sharp decrease can be observed around t_{10}, after which the founding rates of the generalist subpopulation stabilize. Before t_{10} the mean generalist founding rate is 1.6, from t_{10} onward, it is less than half of that (0.78).

Move to the center

According to Assumption 6, decreasing founding rates imply that generalist firms move to the center of the resource base. Such a move should be indicated by increasing competition among generalists. Figure 5.11 shows that firm-on-firm competitive pressure among generalists steadily increases during the whole focal period. Whereas generalists on average targeted 5 to 10% of their competitors' niches in the early years, in 2000 this proportion was around 20%.

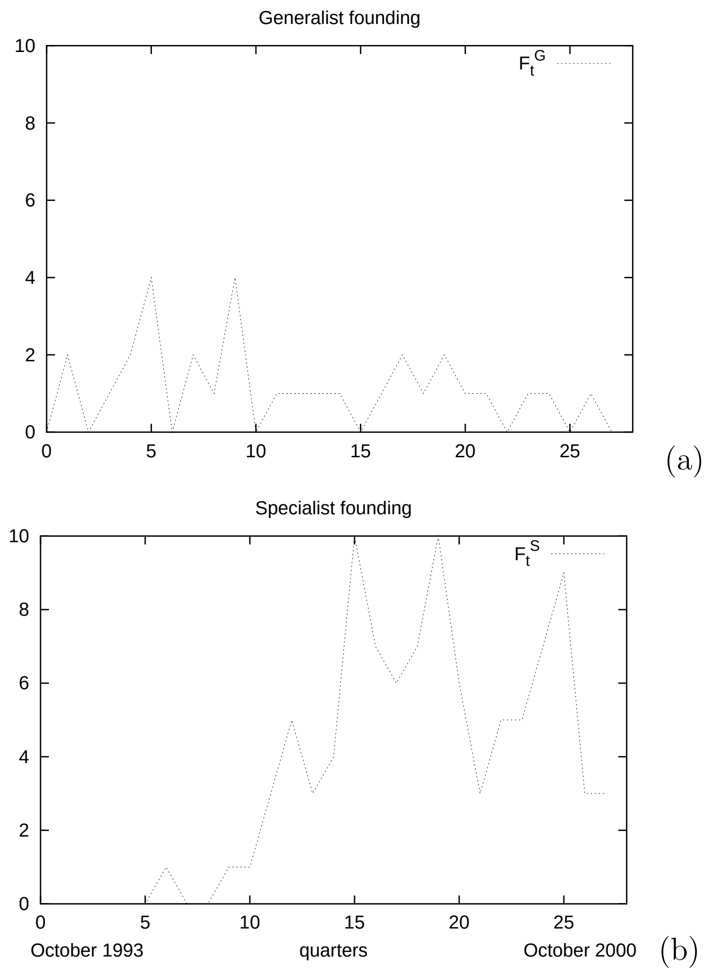


Figure 5.10: Generalist and specialist founding rates

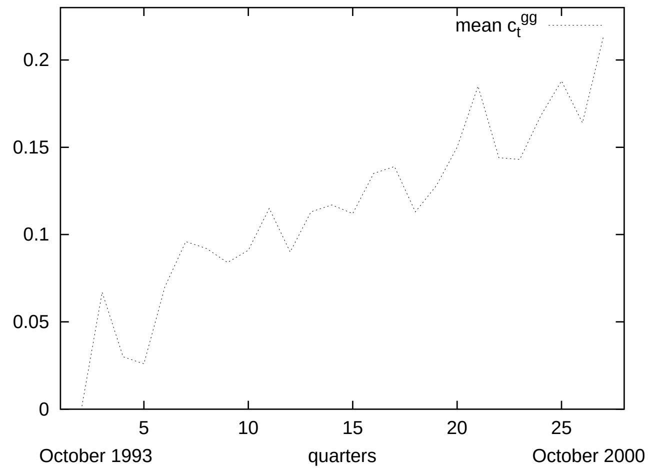


Figure 5.11: Competition among generalists

Resource partitioning

Assumption 7 claims that by moving to the center of the resource base, generalists move out of the periphery. This is not a mathematical tautology: generalists may expand their niches toward the other generalists while at the same time maintaining their positions in the periphery. If generalists move out of the periphery, this should be indicated by a lower competitive pressure imposed by generalists on specialists.

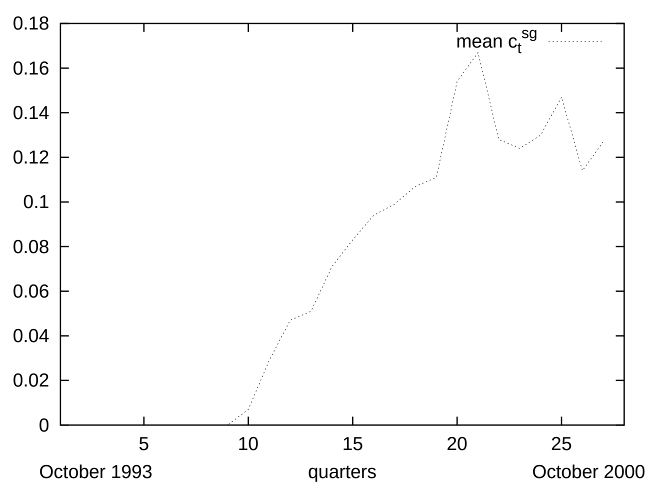


Figure 5.12: Competition imposed by generalists on specialists

Figure 5.12 shows that the competitive pressure by generalists on specialists decreases only from 1999 (t_{21}) onward. Before 1999, the pressure increased, until on average 16% of the niches of specialists was targeted by generalist firms. In 2000, this number has decreased to 12%.

The deceasing firm-on-firm pressure imposed by generalist search engines on specialists suggests that the two subpopulations target—increasingly—different resources, which, in turn, suggests a process of resource partitioning. This suggestion is supported by Figure 5.13, which shows the competitive pressure imposed by specialists on generalists. Here, we see that the mean proportion of generalist

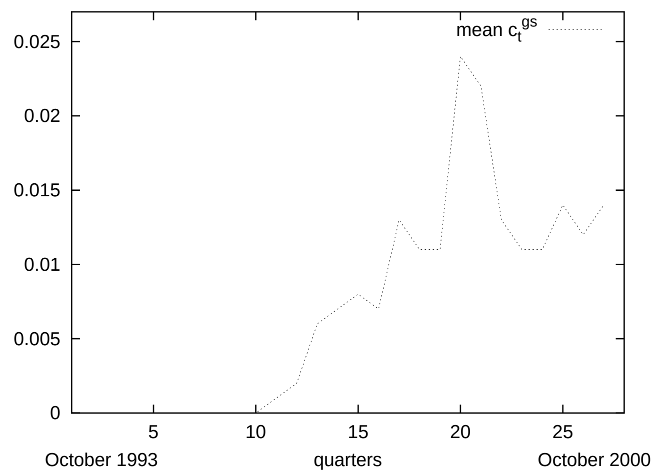


Figure 5.13: Competition imposed by specialists on generalists

niches targeted by specialists decreases from around 2.3% in late 1998 to around 1.3% in 2000.

One may wonder whether this process of resource partitioning is in fact induced by the generalist move to the center. Figure 5.11 shows that generalist search engines have always been moving to the center, whereas resource partitioning began only after 1998. That raises the question if resource partitioning is induced by a generalist move to the center. Possibly, specialists have been moving toward the market periphery, and away from the generalists, and thereby decreased the competitive pressure imposed on them. If specialists move toward the market periphery, they most likely move in different directions. As a consequence, not only the competition between generalists and specialists would decrease, but also the competition among specialists.

Figure 5.14 shows the mean firm-on-firm competition among specialists. It shows that, from the end of 1998 to 2000, the competitive pressure that specialist firms impose on each other reduces by roughly 50%. The implication is, that specialists indeed move toward the market periphery, and consequently, induce resource partitioning.

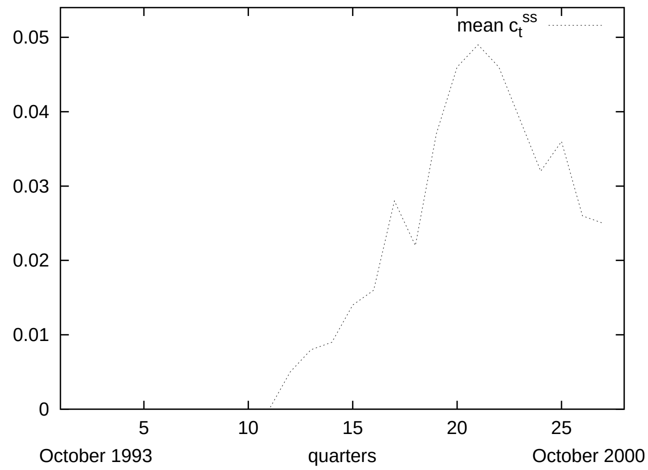


Figure 5.14: Competition among specialists

Specialist competition

Assumption 8 states that the mean competitive pressure on specialist firms decreases if the pressure that generalists impose on specialists decreases. This is not obvious altogether: the mean competitive pressure on specialists is not only exerted by generalists, but also by specialists. As the number of specialists is growing, the competitive pressure on specialists might be growing too.

Figure 5.15 shows that the mean competitive pressure on specialist firms decreases after t_{21}. Whereas before t_{21}, the competitive pressure on specialists in-

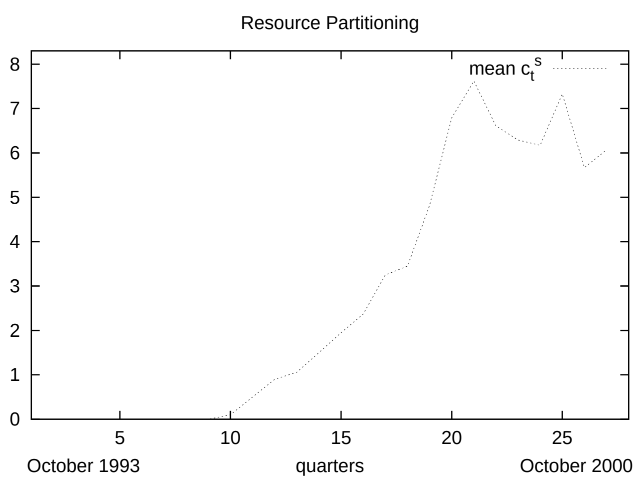


Figure 5.15: Competitive pressure on specialist subpopulation

creases fast, at t_{21} the trend is reversed, and the competitive pressure on specialist firms starts to decrease.

Specialist founding rates

Finally, Assumption 9 states that as a consequence of a decreasing competitive pressure on specialists, the founding rates of specialists increase. Specialist founding rates were shown in Figure 5.10. We see that specialist founding rates observed at later stages are much higher than during the early stages of the market. The suggestion is that specialist founding rates increase. However, as the empirical evidence suggests that resource partitioning processes started only around t_{21}, we cannot relate specialist founding rates before t_{21} to resource partitioning. Figure 5.10 also shows a sudden increase of specialist founding rates after t_{21}. This increase could be related to resource partitioning processes.

So, the empirical evidence suggests that resource partitioning is not the cause for all specialist proliferation in the industry of Internet search engines, but resource partitioning processes, induced by a specialist move to the periphery, are observed, and could have been responsible for specialist inflow at later stages of the market.

In sum, we found some empirical evidence suggesting processes of resource partitioning in the Internet search engine market. We did not find any evidence suggesting processes of density dependence.

5.7 Concentration and dimensionality

In Section 5.3.1 of this chapter we noted that the main claim of resource partitioning theory is that market concentration induces specialist proliferation. In the same section, we mentioned that later versions of the theory contained the claim that an increasing dimensionality of the resource base improves specialist life chances.

We did not test the first claim for theoretical reasons: in Chapter 3, we formalized resource partitioning theory in first-order logic, and it turned out that the theory's main claim did not follow from the theoretical argument. That is, although it may be the case that concentration improves the life chances of specialist firms, resource partitioning theory does not provide the supporting argument.

In the current section, we take a short look at the dynamics of concentration in the industry of Internet search engines, to investigate whether there is empirical evidence for a positive relation between industry concentration and specialist life chances.

5.7.1 Concentration

Our assumptions regarding resource partitioning processes deviate from organizational ecology's original resource partitioning theory, because we replaced each instance of the notion of "concentration" by "economies of scale." We did so, because we concluded from the logical formalization in Chapter 3 that concentration was merely a natural consequence of scale economies; concentration was likely to occur simultaneously with resource partitioning, but would not explain resource partitioning.

The assumption that scale economies imply concentration was part of the formalized theory, though, and now we make it part of our empirical analysis. We have assumed scale economies to dominate in the market of internet search engines, so we expect concentration to increase. Concentration measures are based on the sizes of the firms in an industry. We use niche size as an approximation of firm size, because niche size indicates the number of users of a search engine.

Figure 5.16 shows five different measures of market concentration, based on estimations of the sizes of all search engines in the market. The measures are CR3, CR4 and CR5, the Herfindahl-Hirschman index (Shepherd 1987), and Ijiri and Simon's beta (Ijiri and Simon 1977).[30] In Figure 5.16, all measures are standardized, such that their maximum value is 100. All measures show the

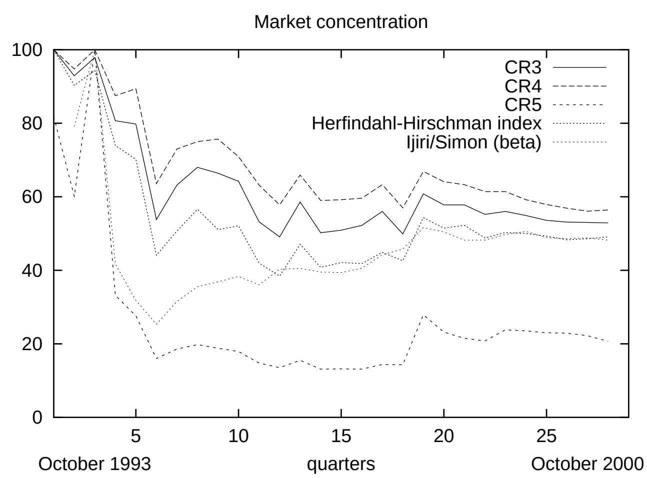


Figure 5.16: Search engine market concentration 1993-2000

same pattern: after an initial decrease of concentration in the first six quarters, concentration stabilizes. Only the Herfindahl-Hirschman index shows a slight increase of concentration from that moment on.

Figure 5.17 shows the log–size for all firms, ranked in size, for the 3rd quarter of every year since 1994. Over the years, the log-size distribution becomes less

[30] In Appendix II, we defined all concentration measures mentioned here.

uneven, and this indicates a more homogeneous distribution of niche sizes.

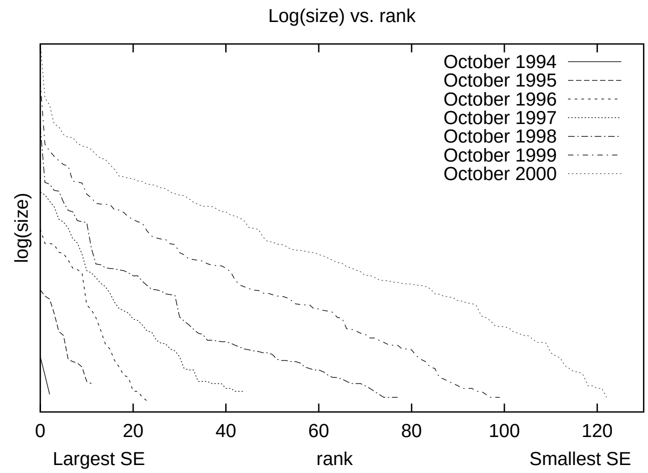


Figure 5.17: Search engine log-size distributions 1994-2000

So, we can conclude that the Internet search engine market did not exhibit an increase in concentration, in spite of its apparent scale economies. The assumption stating that scale economies imply concentration to increase, does not hold in Internet search engine population.

5.7.2 Market dimensionality

In Section 5.3.1, we mentioned that later versions of resource partitioning theory contained a new claim: "As the number of dimensions in the resource space increases, the founding rates of specialist organizations will rise and the mortality rates of specialist organizations will decline" (Carroll and Hannan 2000, p.268).

This claim is adopted from Péli and Nooteboom (1999), and draws upon the *sphere-packing* problem in geometry. This problem involves filling up an d-dimensional (Euclidean) space as efficiently as possible with d-dimensional hyperspheres (*e.g.* balls) of equal size. As it turns out, the efficiency of the densest, *i.e.* optimal, packing decreases as the number of dimensions of the space increases. In other words, as the dimensionality increases, less space is covered by the hyperspheres. This effect is shown to persist up to 20 dimensions; at that point only 0.3% of the space is covered by the hyperspheres.

Péli and Nooteboom suggest that the niches of generalist firms can be represented as d-dimensional hyperspheres of equal size in an d-dimensional resource space, that is shaped by dimensions of consumer taste. As the consumer tastes become more sophisticated, the dimensionality of the resource space increases, and the space that is filled up by generalist niches decreases. Obviously, if the space filled up by the generalist niches decreases, the space complementary to generalist niches increases. In this space, then, firms with smaller niches, *i.e.*, specialists, can thrive.

For the sphere-packing model to apply to organizational niches in resource space, the following assumptions should hold: (1) the niches of all generalists are of equal size, (2) niches do not intersect, (3) niches are distributed over the resource space according to the densest packing, (4) niches are spheres, (5) resources are distributed homogeneously over the resource space, and (6) the resource space is infinite. In the previous sections, we have measured both niche size and niche intersection. It turns out that search engine niches have different sizes. Moreover, search engine niches intersect. The question whether the niches are distributed according to the densest packing cannot be answered, because for intersecting niches packing density is not defined.

In our study we considered the resource *base*, *i.e.*, the set of resources, rather than the resource *space*, *i.e.*, the space where the resources are located. The resource base is an amorphous set, which gives us no information about the shape of the resource space, nor about its distribution and size. The same holds for niches. We considered a niche as a set of resources; the shape of such a set is undetermined. As a consequence, we cannot determine whether either of the assumptions (4) to (6) hold in our empirical domain.

As all assumptions of the sphere-packing approach either do not hold, or are undetermined in the market of Internet search engines, it does not seem to make sense to test for an increasing resource space dimensionality. As a consequence, it does not make much sense to test whether, due to the sphere-packing effect, increasing dimensionality improves the life chances of specialist firms. There is, however, a second argument implying that increasing dimensionality improves the specialist life chances. This argument draws on the so-called "curse of dimensionality."

The curse of dimensionality

The curse of dimensionality is a well known problem in the field of mathematical modeling (Wegman 1990). The "curse" implies that the volume of the periphery of a space grows relative to that of the center of the space, if the dimensionality of the space increases.

Suppose R^* is a d-dimensional space with radius r, and suppose Cen_{R^*}, the center of the space, is smaller than the space itself; that is, the radius of the center of R is $r - \epsilon$ and $\epsilon > 0$. The volumes of R^* and Cen_{R^*} are:[31]

$$V(R^*) = \frac{\pi^{\frac{d}{2}} \cdot r^d}{\Gamma\left(\frac{d}{2}+1\right)} \qquad V(Cen_{R^*}) = \frac{\pi^{\frac{d}{2}} \cdot (r-\epsilon)^d}{\Gamma\left(\frac{d}{2}+1\right)}$$

[31] Notice that: $\Gamma(y) = \int_0^\infty x^{y-1} e^{-x}\, \mathrm{d}x$

The relative volume of center Cen_{R^*} is:

$$\frac{V(Cen_{R^*})}{V(R^*)} = \frac{\pi^{\frac{d}{2}} \cdot (r-\epsilon)^d \cdot \Gamma\left(\frac{d}{2}+1\right)}{\pi^{\frac{d}{2}} \cdot r^d \cdot \Gamma\left(\frac{d}{2}+1\right)} = \frac{(r-\epsilon)^d}{r^d} = \left(1-\frac{\epsilon}{r}\right)^d$$

It holds that:

$$\lim_{d\to\infty} \left(1-\frac{\epsilon}{r}\right)^d = 0$$

So, if $d \to \infty$, the center of the space is empty, and all volume is in the space's edges.

High dimensionality can cause the analysis of multidimensional data to be dominated by edge effects, hence the qualification "curse" (Gershenfeld 1999, 148). For specialist firms, however, the curse of dimensionality may prove to be a blessing in disguise.

Let us assume that the resource space is a space shaped by d taste dimensions. Let us furthermore assume that some tastes are more common than others, that is, resources are not distributed homogeneously over the resource space, but, say, normally. We call the area where most resources are "located" the center of the resource space; the remainder we call the periphery. Resource partitioning theory assumes that generalist firms target the center of the resource space, and specialists target the periphery (Carroll and Hannan 2000, p.268). The curse of dimensionality implies that the proportion of the resource space that lies in the periphery increases if the dimensionality of the space increases. Consequently, resources move into the periphery of the resource base and out of the center; specialist life chances improve, whereas generalist life chances decrease.

Figure 5.18 shows that the sphere-packing approach and the curse of dimensionality approach allocate by and large the same resource space to generalist firms. However, for the curse of dimensionality to apply, less assumptions are obliged to hold.

In the following paragraphs we will try to find out whether the dimensionality of the resource space for search engines has increased in the focal period. Such a finding would give us an additional explanation for the proliferation of specialists in the search engine industry.

Dimensions of taste

We assumed the resource space to be shaped by dimensions of taste. Consumers judge firms on their scores (or "location") on these dimensions. Firms may be considered similar on one dimension, and different on another.

Unfortunately, we do not know the considerations of the consumers in the Internet search engine market, nor do we know the number of dimensions on which they judge search engines, and whether this number changes. We do have measurements of niche intersection, however, which imply that some pairs of

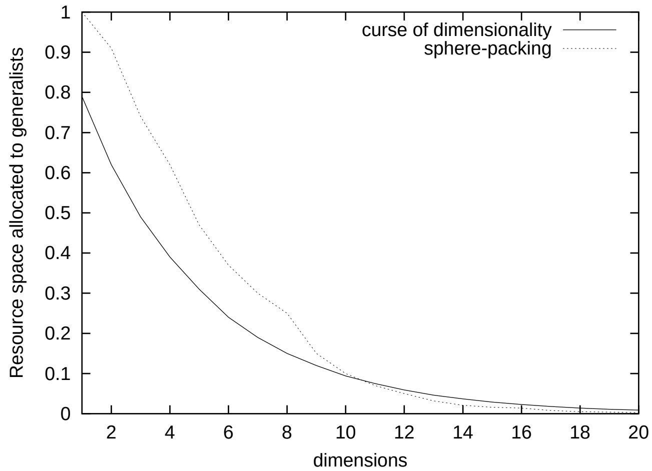


Figure 5.18: Resource space allocated to generalist firms at increasing resource space dimensionality

search engines are perceived similar, and some pairs are perceived different. Our aim is to use these dissimilarities to estimate the dimensionality of the resource base. The reasoning is that the number of taste dimensions should be consistent with the search engine (dis)similarities observed.[32] So, as a first step, we try to determine the (dis)similarity of the niches of the Internet search engines.

Niche (dis)similarities

We assume that two firms that target similar sets of resources, are likely to be perceived as similar by the resources.[33] The more the niches of two search engines intersect, the more similar the search services are perceived to be. However, by random chance, large niches have larger intersections than small niches. To adjust for niche size, we relate niche intersection to the joint niche size of the two firms. In this way we obtain the standard measure for set similarity, Jaccard's J (see, for example, Cox and Cox (1994)). Let A and B be sets; $\cap$ denotes intersection, and $\cup$ the joint size of A and B (disjunction). The Jaccard measure of set similarity is defined as follows:

$$J_{AB} = \frac{A \cap B}{A \cup B}$$

[32] As an example, for two search engines to be perceived dissimilar, one taste dimension is needed. For three search engines to be perceived equally dissimilar, at least two dimensions of taste are needed. To accommodate for many search engines and many firm-to-firm dissimilarities, many dimensions may be needed.

[33] Note that if two firms share a single resource, it is not necessarily the case that the resource perceives the firms as similar. The resource may perceive the firms as substitutable or even complimentary (see also the discussion of Chapter 4). If, however, a number of resources agrees about the simultaneous appeal of two firms, it is likely that they perceive the firms as similar.

In terms of organizational niches, the Jaccard-similarity between the niches of i and j is expressed as:

$$J_{ij} = \frac{v_{ij}}{v_i + v_j - v_{ij}}$$

J_{ij} is between 0 (no intersection) and 1 (full intersection). Using[34] similarity measure J_{ij} we define:

$$d_{ij} = 1 - J_{ij}$$

the dissimilarity, or distance, between the niches of firms i and j.

Determining resource space dimensionality

We use our empirical measurements to calculate d_{ij} for each pair of firms, and thus obtain a symmetrical matrix of niche dissimilarities. The matrix has size $N_t \times N_t$ for each quarter t. Next, we determine the dimensionality of the matrix. One way to determine its dimensionality is by the applying so-called Scree test, or elbow test, a multi-dimensional scaling technique (Steyvers 2002). In a Scree test, all distances in the matrix are represented in a d-dimensional space, as good as possible according to an optimization algorithm. The quality of the representation is determined by a lack of "stress"; stress emerges when distances are represented as either too long, or too short. For a dimensionality d, the stress is calculated. Then the process is repeated for dimensionality $d + 1$, until (1) a either a (low) benchmark value of stress is reached, or (2) the Scree-curve makes a characteristic bend, the so-called "elbow effect". Figure 5.19 shows how stress decreases as the dimensionality of the representation increases, in this particular case for the population of Internet search engines at t_{25}. Appendix III gives a number of illustrations of how the distribution of firms over the resource space can be represented in a 2-dimensional space.

To find out whether the dimensionality of the resource space increases, we calculated the degree of stress associated to 2, 5, 7 and 9-dimensional representations of the distances between all niches in the search engine industry over time. The result is shown in Figure 5.20: as more search engines enter the market, stress increases for all representations.[35] The suggestion is that the dimensionality of the resource base increases.

[34] Although the Jaccard measure adjusts for the fact that, by random chance, large niches have larger intersections than small niches, still, by random chance, pairs of large niches have higher Jaccard similarity than pairs of small niches. To be more precise: by random chance, pairs of niches that are large relative to the resource base, will have higher Jaccard values than pairs of niches that are small relative to the resource base. A measure of set similarity that adjusts for relative niche size, is that of Mozley and Margalev (Cox and Cox 1994), defined as $M_{AB} = \frac{(A \cap B) \times R}{A \times B}$. In terms of niches this amounts to: $M_{ij} = \frac{v_{ij} \cdot R}{v_i \cdot v_j}$. An important drawback of similarity measure M is that it is not metric, which complicates the estimation of the dimensionality of the resource base.

[35] The multi-dimensional scaling tool applied here was UCINET. We only performed the analyses for t_0, t_5, t_{10}, t_{15}, and t_{25}; hence the stylized curves.

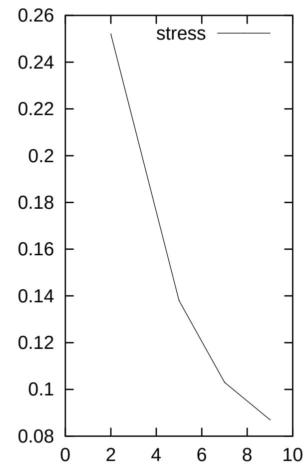


Figure 5.19: Stress per dimension at t_{25}

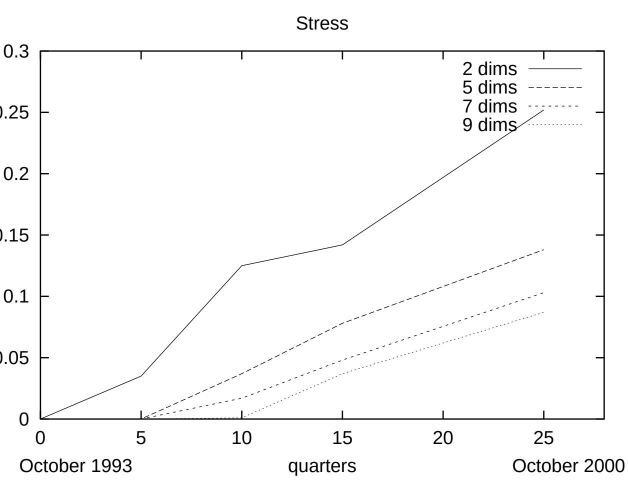


Figure 5.20: Stress associated with scaling at several dimensionalities, over time

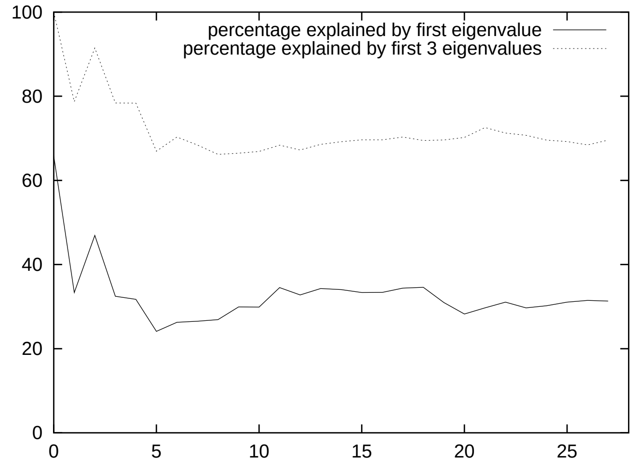


Figure 5.21: Eigenvalue ratios

Eigenvalues and the distribution of resources

The curse of dimensionality implies that if the dimensionality of the resource space increases, the portion of resource space available to specialists increases. Firms, however, do not thrive on resource *space*, but on resources. So, before we accept increasing dimensionality as an explanation of specialist proliferation, we should test whether specialists actually make use of the newly available space.

To do that, we use metric MDS to calculate the *eigenvalues* of the dissimilarities-matrix D. An eigenvalue indicates the relative importance of a dimension.[36] If many specialists make use of peripheral space that is induced by growing dimensionality, the eigenvalues associated with higher dimensions should increase with respect to the eigenvalues associated with lower dimensions (Cox and Cox 1994).

Figure 5.21 displays the proportion that the first and the first three eigenvalues constitute of the sum of the first 5 eigenvalues. As it turns out, the first eigenvalue accounts for a stable 30% of the total, whereas the first three eigenvalues account for a stable 70%. The implication is that higher dimensions, *i.e.*, dimensions 4 and 5, do not gain importance: they explain a stable, but limited portion of the dissimilarities found between the niches of Internet search engines.

In sum, increasing dimensionality does not induce the life chances of specialists in the Internet search engine industry. Although we found evidence that the number of dimensions of the resource base has increased over the focal period, and although in the same period the specialist life chances have improved, the two observations are not related. The analysis of the eigenvalues suggests that

[36]Eigenvalues give the relative weight of eigenvectors. The first eigenvector corresponds to the most important dimension, the second eigenvector the second most important dimension, etc.

the space generated by increasing dimensionality is, for the greater part, not used by specialist search engines.

5.8 Discussion and conclusions

We started this chapter with the observation that in a particular industry, *i.e.*, Internet search engines, both density dependent and resource partitioning processes appeared to occur. This observation immediately raised a theoretical issue, as both processes are not supposed to occur simultaneously. Density dependence theory describes how a population of firms emerges and reaches an equilibrium state, in which organizational founding and disbanding rates balance each other out. The theory's explanatory mechanism applies from the moment that the first firms have entered the market, until the moment that the equilibrium is reached. Resource partitioning theory describes how the equilibrium can be disturbed when a partitioning of resources causes a proliferation of specialist firms.

In the industry of Internet search engines, specialist firms start to proliferate before an equilibrium between founding and mortality rates is reached. The question is whether the specialist proliferation can be explained by resource partitioning theory.

In order to be able to identify resource partitioning processes in an emerging industry, we developed a mathematical model for resource partitioning theory build on, and consistent with, the model of density dependence theory. The new model has mathematical counterparts for qualitative phrases such as "generalists move to the center of the resource base," and "the market has become partitioned into generalist and specialist resources." Furthermore, the model is able to isolate the processes relevant for specialist proliferation, regardless whether an industry is in an equilibrium state or not.

The key to connecting density dependence theory and resource partitioning theory lies in the refinement of the concept of competition that was discussed in Chapter 4. In Section 5.3.3 we introduced the refined concept of competition in density dependence theory, and showed that the predictions of density dependence theory did not change.

The conceptual refinement of competition made it a variable that could be independently measured. In organizational ecology's original density dependence theory, competition was defined on population size.

We altered the other key concept of density dependence theory, legitimation, as well, and made a measurable variable. In the original density dependence theory, legitimation, too, was defined on population size.

By making competition and legitimation independently measurable variables, we have induced the explanatory power of density dependence theory. Because competition and legitimation can be measured independently, the assumptions that specify the relations between population size and both competition and legitimation are falsifiable. In organizational ecology's original density dependence theory, both assumptions were not falsifiable. By adding falsifiable assumptions a theory, the explanatory power of the theory is enhanced (Popper 1959).

The fact that both added assumptions were falsified by the empirical domain does not make the initial theory less powerful; it merely clarifies that in the case of Internet search engines something special is going on: competitive pressure is lower than predicted, and legitimation is higher than predicted due to—as evidence suggests—differentiation in the industry.

The two assumptions stating that competition and legitimation determine vital rates were also falsified. Let us review some possible reasons why the relations specified in both assumptions were not observed.

First, it could be that our refinement of competition and our alternative definition of legitimation are erroneous. We aimed at measures that captured our intuitions about competition and legitimation, but these intuitions could have been wrong or incomplete.

A second reason for competition and legitimation's failure to predict founding and mortality rates could be in the collection of the data. We assumed that the number of references to a particular search engine reflects the number of users—or at least potential users—of the engine. By doing so, we assumed that the attitudes and preferences of Internet users who have an Internet page and mention search engine URLs on their page, are representative for the attitudes and preferences of all Internet users. This may not be true. Although the sample of Internet pages that we have taken is large (many millions), it may have been biased toward pages of Internet professionals or devotees.

Third, it could be that our period of measurement is too short to observe the expected relations. Most empirical studies of density dependent processes cover a number of decades, rather than a number of years. Expecting competition and legitimation levels to have observable impact within a few months may be naïve, especially in a market subject to many possible intervening variables. The unparalleled economical conjuncture, the disproportionate public attention, and the enormous expectations about Internet's possibilities may have had more influence on vital rates than competitive pressure or legitimation had. Our study focused on an industry over a short, and eventful, period. As a consequence, intervening variables may have dominated the results.

Fourth, it may be the case that legitimation and competition in fact do not predict founding and mortality rates. Organizational ecology's original density dependence theory contained only two falsifiable claims: industrial founding rates are $\cap$-shaped over time and mortality rates are $\cup$-shaped. In the search engine industry, founding rates turn out to have a $\cap$-shaped trajectory, but, as it seems, without competition and legitimation having anything to do with it.

In our view, the lack of predictive power of both legitimation and competition suggests that their effects on organizational vital rates are weaker than assumed by density dependence theory. However, to support our view, the effects of legitimation and competition should be studied over a much longer, and maybe a less turbulent, period of observation.

Our evidence for the claims of resource partitioning theory is, admittedly, somewhat circumstantial. Most assumptions in resource partitioning theory have a rather loose "if A occurs, then B occurs" structure, for which empirical evidence is not easily obtained. An assumption stating "if A occurs, then B occurs" suggests that B should happen after A happens, or at least not before A happens. But a time frame within which A should lead to B, or an expected duration of either A or B are not specified. As a consequence, looking for evidence for resource partitioning's assumptions may turn into a somewhat subjective affair.

Nevertheless, we feel that we have found convincing evidence for "proper" resource partitioning, *i.e.*, the observation of a decreasing intersection between generalist niches and specialist niches. Before 1999, the degree in which generalist and specialist firms appealed to the same consumers was increasing rapidly. In the beginning of 1999, the process turned around, and the intersection started to decrease. This had a clear impact on the competitive pressure imposed on specialists. For this variable, too, the rapid increase turned around, even though specialist competitors were still entering the market *en masse*.

Resource partitioning theory predicted specialist founding rates to increase after March 1999. This suggests that the peak in specialist founding rates that was observed from June 1999 to March 2000 could be caused by resource partitioning processes.

From the logical formalization of resource partitioning theory in Chapter 3 of this thesis, we concluded that market concentration had no explanatory role in the argument. We considered concentration, along with resource partitioning processes, to be implied by scale economies. In Section 5.7, we showed that in the case of the Internet search engine industry, scale economies do *not* imply concentration. This finding could be due to the fact that the industry is growing. If the size of an industry remains stable, than concentration inevitably increases if scale economies dominate (Schmalensee and Willig 1989). However, in a theory of resource partitioning that refers to growing populations as well, the assumption that scale economies imply concentration should be dropped. It does not add any explanatory value, and, in growing industries, it may not hold.

In Section 5.7, we also investigated resource partitioning theory's recent claim stating that an increasing dimensionality of the resource base causes specialist life chances to improve. We dismissed the original supporting argument—the sphere packing model—because it did not apply to the industry of Internet search engines. None of its assumptions about niche sizes and intersections were found to be true. As an alternative supporting argument we considered the "curse of dimensionality." We showed (Figure 5.18) that the curse of dimensionality and the sphere packing model have by and large the same implications. Figure 5.18 also shows the limitations of the sphere packing model; if the model applies, then the curse of dimensionality also applies, ensuring that the center of the resource space becomes empty as its dimensionality increases. The implication is, that whereas the sphere packing model increases the space between the generalist niches, the curse of dimensionality empties this space at approximately the same rate. An argument in favor of the sphere packing approach is that it, contrary to the curse of dimensionality, applies to infinite spaces. However, resources are not infinite, and an infinite resource space that contains a finite number of resources is similar to a diluted solution; the probability for a finite hypersphere to contain a resource is 0. The implication is that in an infinite resource space, niches are empty.

In the introduction of this chapter, we have asked the question whether the differentiation observed in the Internet search engine industry was due to the lenient character of the market, or to established firms forcing new firms into specialism. On the one hand, we have found almost no cases of disbanding firms. This suggests that competitive pressure may have been low. On the other hand, we observed a specialist move to the market periphery, which suggests that competitive pressure may have been something to reckon with.

The fact that we have observed nearly no disbanding search engines is not necessarily due to a lack of competition. First, the observation could be an artifact of the data collection. If a search engine that is referred to on a number of Internet pages is discontinued, the references will remain where they are; in many cases for years. As a consequence, references to search engines may be observed years after the services were discontinued.

A second explanation for the fact that we hardly observed disbanding, can be found in a peculiar feature of the search engine industry, namely, in a lack of need to exit the market. We have encountered a number of search engines that provided a search service, but appeared to be no longer maintained. Apparently, there was no need for the search engines to go through the process of actual discontinuation, even after the organization maintaining the search engine was dismantled.[37]

In sum, in order to investigate differentiation processes in the emerging market of Internet search engines, we developed a mathematical model of resource partitioning theory, based on the model of density dependence theory. In order to

[37] An example of a unmaintained search engine was `http://www.altavista.magallanes.net/`, a search service aiming at the south-American continent. When we consulted the service in the beginning of 2001, it turned out that it had indexed around 300,000,000 Internet pages dated before March 1999, and around 300 pages dated after March 1999. That is, almost all results that the search engines came up with were over two years old, a majority of them dead links.

connect both organizational ecological theories, we used a measure of competition that we developed in Chapter 4. In the course of modeling resource partitioning theory, we also improved the explanatory power of density dependence theory. We measured pairwise competition for 137 members of the Internet search engine industry, in the first seven years of its emergence. We found evidence for resource partitioning: small firms decrease their niche intersection with larger competitors by moving into the market periphery. We found no evidence for density dependence: the relations between competition, legitimation and population size that were proposed by density dependence theory, were not observed.

Appendix I: Proofs of Theorems 1 and 2

We prove Theorems 1 and 2 by propositional logic. First, we introduce propositional logic. Second, we formalize our assumptions and theorems in propositional logic, and third, we prove the theorems. Readers who are familiar with propositional logic may want to skip the introduction and proceed to the formalization.

Propositional logic

Symbols. Propositional logic has the following symbols:

- propositions: p_1, p_2, etc.
- connectives: $\wedge$ (and), $\vee$ (or), $\rightarrow$ (if ... then ...)
- negation: $\neg$
- auxiliary symbols, such as (and)

Formulas. The symbols can be used to make formulas. A single proposition is a formula. If p_1 and p_2 are formulas, then:

- $\neg p_1$
- $(p_1 \wedge p_2)$
- $(p_1 \vee p_2)$
- $(p_1 \rightarrow p_2)$

are also formulas.

Truth values. Each formula has a truth value.

- $\neg p_1$ is true iff p_1 is false
- $(p_1 \wedge p_2)$ is true iff both p_1 and p_2 are true

- $(p_1 \vee p_2)$ is true iff either p_1 or p_2 is true
- $(p_1 \rightarrow p_2)$ is true iff p_2 is true or both p_1 and p_2 are false

Derivation Rules. From formulas we can derive other formulas, by use of derivation rules. Two of these derivation rules are:

- R1: If $(p_1 \rightarrow p_2)$ and $(p_2 \rightarrow p_3)$, then $(p_1 \rightarrow p_3)$
- R2: If $(p_1 \rightarrow p_2)$ and $(p_1 \rightarrow p_3)$, then $(p_1 \rightarrow (p_2 \wedge p_3))$

Formalization

Our assumptions are the following:

- A5: If scale economies dominate, then $N_t''^G < 0$
- A6: If $N_t''^G < 0$, then $\bar{c}_t'^{gg} > 0$
- A7: If $\bar{c}_t'^{gg} > 0$, then $\bar{c}_t'^{sg} < 0$
- A8: If $\bar{c}_t'^{sg} < 0$, then $\bar{c}_t'^S < 0$
- A9: If $\bar{c}_t'^S < 0$, then $N_t''^S > 0$

Below are the theorems we aim to prove:

- TH1: If scale economies dominate, then $\bar{c}_t'^{gg} > 0$ and $\bar{c}_t'^{sg} < 0$
- TH2: If scale economies dominate, then $N_t''^G < 0$ and $N_t''^S > 0$

We use p_1 to denote "scale economies dominate," p_2 to denote $N_t''^G < 0$, p_3 to denote $\bar{c}_t'^{gg} > 0$, p_4 to denote $\bar{c}_t'^{sg} < 0$, p_5 to denote $\bar{c}_t'^S < 0$, and p_6 to denote $N_t''^S > 0$. Our assumptions and theorems can be formalized as follows:

- A5: $p_1 \rightarrow p_2$
- A6: $p_2 \rightarrow p_3$
- A7: $p_3 \rightarrow p_4$
- A8: $p_4 \rightarrow p_5$
- A9: $p_5 \rightarrow p_6$
- TH1: $p_1 \rightarrow (p_3 \wedge p_4)$
- TH2: $p_1 \rightarrow (p_2 \wedge p_6)$

Proof

Below, we prove both theorems. We use Li to denote intermediate results (lemma's).
Proof of TH1:

- From A5 and A6 we derive, by R1:
  L1: $p_1 \rightarrow p_3$
- From L1 and A7 we derive, by R1:
  L2: $p_1 \rightarrow p_4$
- From L1 and L2 we derive, by R2:
  TH1: $p_1 \rightarrow (p_3 \wedge p_4)$.

Proof of TH2:

- From L2 and A8 we derive, by R1:
  L3: $p_1 \rightarrow p_5$
- From L3 and A9 we derive, by R1:
  L4: $p_1 \rightarrow p_6$
- From A5 and L4 we derive, by R2:
  TH2: $p_1 \rightarrow (p_2 \wedge p_6)$.

Now, both theorems are proven.

Appendix II: Five measures of market concentration

Below are the definitions of five measures of concentration that we applied in this chapter.

C3. Let i_1, i_2 and i_3 be the largest three firms in the industry. Let s_i be the size of firm i, and let N be the number of firms in the industry. Then:

$$C3 = \frac{s_{i_1} + s_{i_2} + s_{i_3}}{\sum_{i=1}^{N} s_i}$$

C4 and C5. Let i_1, i_2, i_3, i_4 and i_5 be the largest five firms in the industry. Then:

$$C4 = \frac{s_{i_1} + s_{i_2} + s_{i_3} + s_{i_4}}{\sum_{i=1}^{N} s_i}$$

Similarly:

$$C5 = \frac{s_{i_1} + s_{i_2} + s_{i_3} + s_{i_4} + s_{i_5}}{\sum_{i=1}^{N} s_i}$$

Herfindahl-Hirschman index, HHI. Let $s_{TOT} = \sum_{i=1}^{n} s_i$. Then:

$$HHI = \sum_{i=1}^{n} \left(\frac{s_i}{s_{TOT}} \cdot 100 \right)^2$$

Ijiri-Simon's β. Let r_i be the rank of firm i in a size ordering. Let M be the largest firm in the industry, then:

$$\beta = \frac{\sum_{i=1}^{N} \frac{log\left(\frac{M}{s_i}\right)}{log(r_i)}}{N}$$

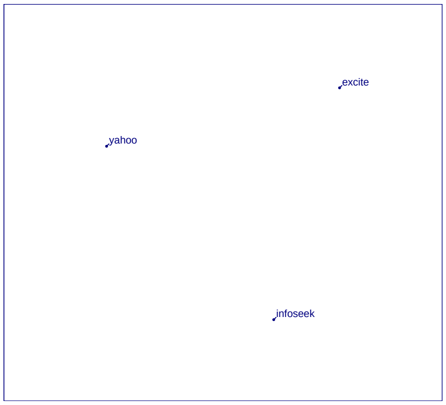
Appendix III: 2-dimensional representations of the search engine industry in resource space

In Figure 5.22 we optimized the relative positions of a subset of the search engine population in a 2-dimensional space,[38] for 6 out of the 28 available quarters. In the beginning of 1994 (Figure 5.22(a)) there are only three search engines around, and their mutual distances can be represented in a 2-dimensional space without stress. As the subpopulation grows, the stress for the 2-dimensional representation increases. In particular at later times, search engines that have the same domain (.nl, .de, .au, .ru[39]) are perceived similar by the consumers: they are located close together. Moreover, the center part of the space is populated by large generalist search engines, whereas the periphery is populated by smaller specialists.

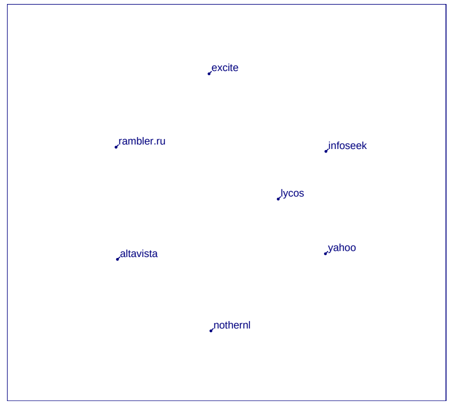
Figure 5.23 shows the optimal two-dimensional representation of the relative locations of the complete search engines population in the second quarter of the year 2000. In this figure 134 search engines are considered, that is, all search engines with a non-empty niche. The stress generated by 134 search engines is higher than the stress generated by the subpopulation of 21 search engines as represented in Figure 5.22(f). If we look at the picture closely, we again see the proximity of search engines originating from the same, or similar, domains.

[38]Note that we have not attempted to determine what the dimensions signify. We chose a 2-dimensional representation of the search engine industry in resource space not because we were interested the particular locations of the engines on the dimensions, but because a 2-dimensional representation can be conveniently displayed.

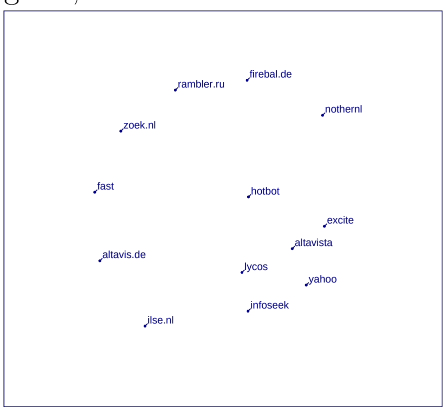
[39]The domains .nl, .de, .au, and .ru indicate that these search engines originate from, respectively, The Netherlands, Germany, Australia and Russia.



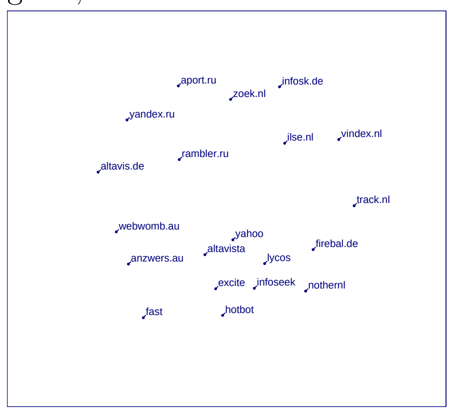
(a) quarter 1994/1; 3 search engines; stress 0.



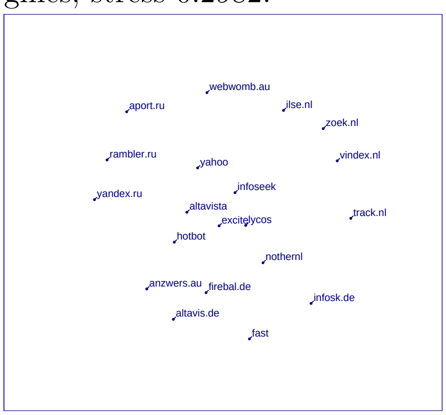
(b) quarter 1995/2; 7 search engines; stress 0.2512.



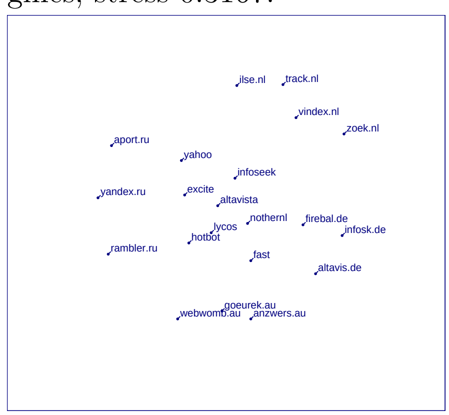
(c) quarter 1996/3; 13 search engines; stress 0.2982.



(d) quarter 1997/4; 20 search engines; stress 0.3167.



(e) quarter 1999/1; 20 search engines; stress 0.2825.



(f) quarter 2000/2; 21 search engines; stress 0.2646.

Figure 5.22: Two dimensional plots of six quarters (metric MDS on 2 dimensions).

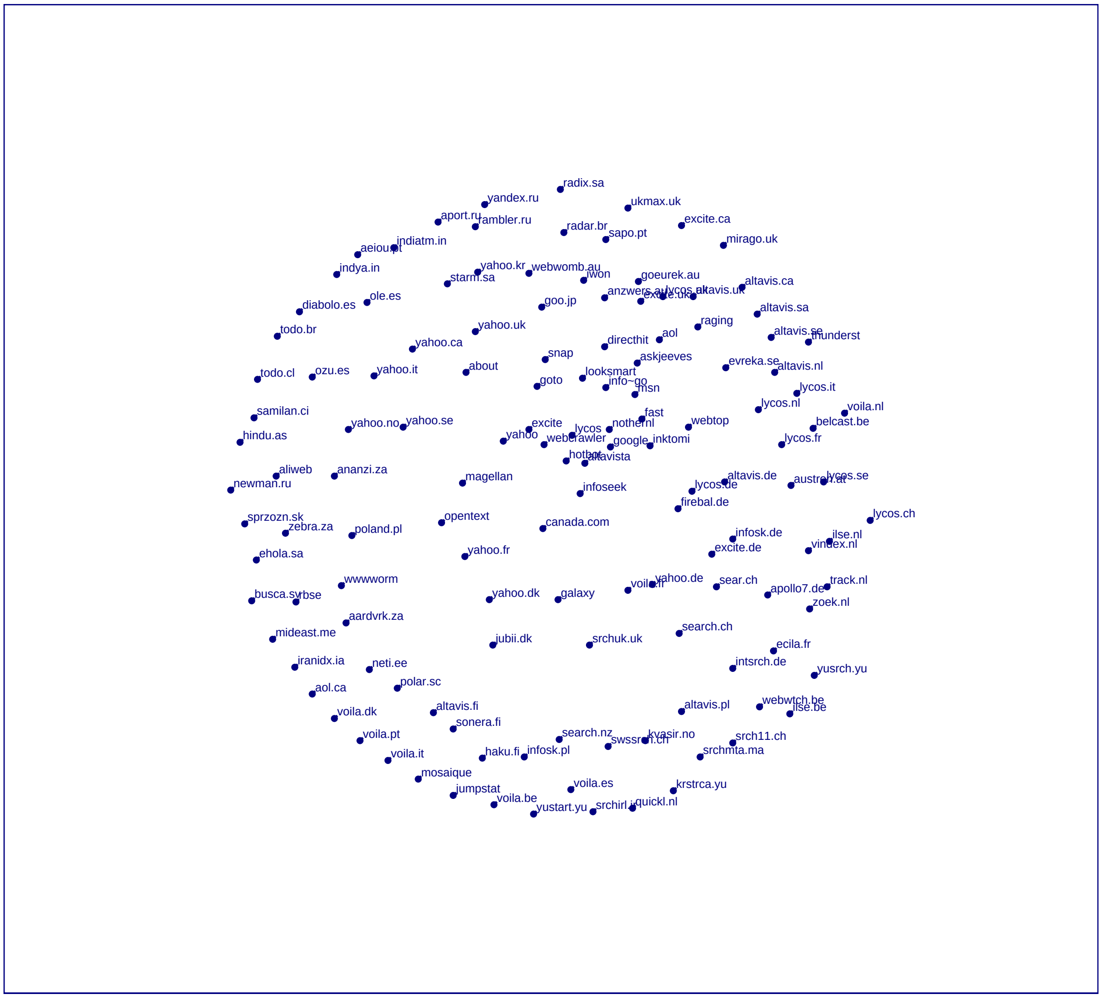


Figure 5.23: Two dimensional plots of quarter 2000/2; 134 search engines; metric MDS on 2 dimensions; stress 0.3674.

Chapter 6

Results and Conclusions

The aims of this thesis are, in the most general terms:

1. Developing heuristics for theory formalization.
2. Applying formalization to social science theories.
3. Clarifying social processes.

In Chapter 2 we developed a step-by-step approach to logical theory formalization. In Chapter 3 we applied our approach to a social science theory: resource partitioning. In Chapter 4, we used social network techniques to analyze the concept of competition in industries, and to formalize a number of notions about competition. And in Chapter 5, we used the formalizations of Chapter 3 and 4 to connect resource partitioning theory to density dependence theory. As a consequence, we were able to clarify the competitive dynamics in the market of Internet search engines.

In this concluding chapter, we first present the results of the individual chapters. Subsequently, we discuss the results with respect to our three aims. We end with a conclusion.

Chapter 2. Developing a systematic approach to theory formalization

We developed a 5-step approach to computer supported logical formalization of (social) science theories. The approach takes a text containing an scientific argument as a point of departure, and helps to produce a formal, sound, and consistent theory as a point of termination. The approach was designed to target formalization systematically. Each subsequent step has the output of its preceding step as an input, and for each step, a number of heuristics is presented that help to gain insight in a theory.

The first three steps of our approach, constituting a so-called rational reconstruction, focus on reducing logical and conceptual ambiguity. Step 1 identifies sentences in the text that capture the core theory. This is done by first focusing on the main claims, and subsequently on their supportive arguments. The output of step 1 is a list of sentences quoted literally from the text. Step 2 analyzes and sharpens key concepts in the core theory. The output of step 2 is a "dictionary" of key concepts. The dictionary increases the parsimony of the theory by relating concepts to each other. Step 3 analyzes the line of argumentation. The goal is to represent the core theory as a set of relatively simple sentences, with a clear logical structure. In addition, for each sentence, its role in the argument is specified: premise, or conclusion.

The first three steps constitute the rational reconstruction, which we consider to be the most important part of a formalization effort. The rational reconstruction clarifies and disambiguates theories, and explicates the understanding that the formalizer has of the theories, thereby facilitating the scientific debate.

The next two steps constitute the actual formalization: in step 4, the set of sentences that result from the rational reconstruction is represented in formal logic. When all statements of the core theory are represented formally, in step 5, attempts are made to prove the theorem candidates. Some steps may have to be repeated: formalization is an iterative process.

To illustrate the merits of our approach, we used an example from resource partitioning theory. Using the example, we went through the formalization approach step by step, and made our choices explicit.

Chapter 3. Formalizing resource partitioning theory

By use of the 5-step formalization approach that was presented in Chapter 2, we systematically thought through resource partitioning theory, a fragment of organizational ecology. We took a scientific text as a point of departure. In the first step of our formalization approach, we identified seven sentences that constituted the core theory of resource partitioning. In the second step we analyzed ten important concepts that occurred in the seven sentences. We determined the meaning of the concepts and their mutual relations.

In the third step, we "informally axiomatized" the seven sentences. The result was a set of eight informal axioms labeled assumptions (premises), and two informal labeled theorems (conclusions). Of each informal axiom, the logical structure was clarified and the meaning of its concepts determined. Furthermore, we—informally—derived four lemmas (intermediate conclusions), identified an implicit background assumption, and composed a conceptual model of the explanatory argument.

In the fourth step, we formalized the informal axioms in first-order logic. All resulting formal axioms turned out to be universally quantified and to have simple logical structures. In the fifth and last step, we tested whether the formal theorems could be proved from the assumptions. To this end, we used an automated theorem prover. It turned out that we needed to formulate an additional background assumption, as well as four corollaries, to be able to prove both theorems.

By adding an assumption to the theory, we were able to derive a new theorem that explains a social process that had been empirically observed, but had previously only been tentatively related to resource partitioning processes.

The main result of our formalization is a sound and consistent logical representation of resource partitioning theory. An additional result of the formalization is the increased parsimony of resource partitioning theory. At least six assertions—suggested by several theorists to help explain resource partitioning—were proved to be redundant, as was some of the information in the core theory.

Chapter 4. Competition in industries

In Chapter 4 we made a conceptual analysis of "competition" in the context of organizational ecology, and of the concept's relation with other concepts. This effort can be regarded as an extended version of the conceptual analysis that is the second step in our 5-step formalization approach.

The result of the conceptual analysis is a network-representation of competition in industries that formally connects competition at the micro level of individual firms to competition at the macro level of industries. The micro-macro connection facilitates the connection between theories *about* the micro level of competition—such as resource partitioning theory, which is discussed in Chapter 3—and theories about the macro level of competition—such as density dependence theory, which is discussed in Chapter 5.

Network analysis was chosen as a means of conceptual analysis because it is a convenient way to analyze relations in groups, and because it provides aggregation techniques that facilitate establishing a micro-macro link.

We started out with a basic—bimodal—network-representation of a market, consisting of the appeals that firms make to resources. Upon the market representation, we defined a—unimodal—network-representation of competitive relations between firms.

We made a distinction between (1) the competitive relation itself, (2) the intensity of the competitive relation, and (3) the pressure that the relation imposes on both competitors. Our intuition was that two firms entertain a competitive relation if they appeal to the same resource. The intensity of a pairwise competitive relation depends on the number of resources that two firms share. The competitive pressure imposed by firm j on firm i is determined by the proportion of i's resources to which j appeals. To make the intuitive notions formal, we defined a dichotomous measure of a competitive relation, a continuous measure of competitive intensity, and a continuous asymmetric measure of competitive pressure.

By comparing the actual number of competitive relations in an industry to the potential number of competitive relations, we obtained a measure of the density of the competitive network. By aggregating the competitive pressure imposed on one firm by all other firms in the industry, we obtained a measure of total competitive pressure on a firm. By aggregating the competitive pressure on all firms in an industry, we obtained a measure for industry competition, thereby establishing the micro-macro link in competition. We showed that under *ceteris paribus* conditions, industry competition will increase with the number of firms in an industry at an increasing rate, as is assumed by organizational ecology.

Upon the basic market representation, we defined a notion of organizational niche, and showed that our measure of competitive intensity formalizes the crucial ecological notion of niche intersection. We investigated the relation between niche size and competition. It turned out that niche size is an important factor in pairwise competition: firms with large niches impose more competitive pressure on firms with small niches than vice versa. The relation between niche size and industry competition is less straightforward: firms with large niches impose more competitive pressure than do firms with small niches, but receive by and large the same competitive pressure. This finding is surprising, as it appears to oppose Barnett's (1990) notion that large firms are weak competitors and strong survivors.

We measured competition at the micro level in the industry of Internet search engines over a seven year period. By means of the micro-macro link, we were able to test a number of conjectures about competitive dynamics at the macro-level.

Chapter 5. Competition in the industry of Internet search engines

Chapter 5 empirically investigated the emergence of the industry of Internet search engines, from the point of view of organizational ecology. The industry of Internet search engines dates back to 1993. From 1997 onward, the industry has shown a proliferation of specialist search engines that offer services to particular geographical domains or language areas.

To understand the success of the specialist subpopulation in the emerging Internet search engine market, we set out to connect resource partitioning theory, which explains specialist proliferation, with density dependence theory, which explains industry emergence. Establishing the connection is complicated by the fact that (1) resource partitioning theory assumes competitive pressure to differ for different firms in an industry, whereas density dependence theory employs one measure of competition for the entire industry. The fact that (2) resource partitioning theory applies to mature industries, whereas density dependence theory applies to emerging industries, complicated a possible connection between both theories as well.

By means of the micro-macro link in competition that we established in Chapter 4, we found a solution for the first problem. The solution for the second problem lay in restating resource partitioning theory by use of relative, rather than absolute measures of competition and vital rates.

We started out by replacing the measure of industry competition in density dependence theory by the measure we developed in Chapter 4. The replacement had two advantages. First, it facilitated a connection with resource partitioning theory. Second, our notion of industry competition can be measured independently of population size, which adds a falsifiable assumption to density dependence theory, and enhances the theory's explanatory power. We altered density dependence's measure of legitimation in a similar fashion, again enhancing density dependence's explanatory power.

Build on the logical analysis of Chapter 3, and the conceptual analyses of Chapter 4, we formulated a theory of resource partitioning that applies to emerging industries.

We measured the competitive dynamics in the market of Internet search engines and tested a number of assumptions. It turned out that, contrary to the assumptions made by density dependence theory, competition in the search engine industry does not increase at an increasing rate. Instead, it follows an S-shaped trajectory. Legitimation does not increase at an decreasing rate, as assumed by density dependence theory, but at an increasing rate. The deviant trajectories of both competition and legitimation are due to a decreasing pairwise competitive pressure in the search engine industry, which was also observed in Chapter 4. Moreover, legitimation and competition failed to predict founding and mortality rates, as was assumed by density dependence theory.

Whereas all assumptions constituting density dependence theory were falsified, most of the assumptions constituting resource partitioning theory were confirmed. In the Internet search engine market, the growth rate of the generalist subpopulation decreases, and the competition between generalists increases, indicating a generalist move to the center of the market. Resource partitioning theory assumes this move to result in a decreasing competitive pressure of generalists on specialists and vice versa, indicating resource partitioning "proper." Although the partitioning of resources is clearly observed in the search engine market, evidence suggest that it may be a consequence of specialists moving toward the market periphery, rather than of generalists moving toward the market center. Because of resource partitioning, the competitive pressure on the specialist subpopulation decreases, which, in turn, improves specialist life chances. Indeed, in 1999 and 2000, we see a peak in specialist inflow. Specialist inflow before 1999, however, cannot be attributed to resource partitioning, and thus needs to be explained in a different way.

A possible explanation for the early proliferation of specialist firms is Péli and Nooteboom's (1999) sphere-packing model, which is incorporated in later versions of resource partitioning theory. The model explains how an increasing dimensionality of the resource space improves the life chances of specialists. A number of assumptions of the model, however, do not hold in the Internet search engine market. Consequently, the model does not apply, and cannot explain specialist proliferation. As an alternative, we showed that by means of the so-called "curse of dimensionality," increasing resource space dimensionality may improve specialist life chances; contrary to the sphere-packing model, the curse of dimensionality does apply to the Internet search engine market. We define a measure of niche (dis)similarity, and use the measure to determine the dimensionality of the Internet search engine resource space, per quarterly assessment. Evidence suggests that the dimensionality of the search engine resource space increases in the focal period. However, we found no evidence that specialist search engines thrive on resources that are made available by higher dimensions. An analysis of eigenvalues showed that a stable 70% of the dissimilarities between search engine niches is explained by the first three dimensions.

This concludes the results of Chapters 2 to 5 of this thesis. We proceed by discussing the results with respect to the three aims that we set in the introduction of this thesis.

Aim 1. Developing heuristics for theory formalization

We developed a systematic approach to (logical) formalization of social science theories that may help social scientists to produce formal representations of theories in their field. We intentionally presented our formalization approach in an informal manner. By doing this, we aimed to address social science theorists, rather than logicians or computer scientists.

We also intentionally referred to our approach as "heuristics that help to formalize" instead of a "method that results in a formalization." Clearly, our approach is not deterministic; different formalizers will come up with different formalizations of the same scientific text. Formalizers may have differences of opinion as to what parts of the text constitute the relevant scientific argument, what the concepts mean, and what is the logical structure of the argument presented. Even the translation of informal axioms into formal axioms may differ from person to person.

Therefore, we never meant our formalization approach to "objectify" the meaning and the implications of theories. The interpretation of—in particularly—social science theories will always be subject to personal interest and preference. We meant our formalization approach to explicate the choices that interpretors make in the course of understanding a theory. Our formalization approach makes a number of important choices in theory understanding explicit, tractable, and retractable, thereby facilitating the scientific debate, not only with colleagues, but with oneself as well.

An important advantage of the step-by-step character of our formalization approach is that it controls the explosion of ambiguity that comes with informal theory interpretation. We showed that a simple, and perfectly reasonable, argument in natural language, consisting of seven sentences, may easily be interpreted in billions of different ways. By going through the interpretation process step by step, we usually have to consider only a few possible interpretations at the same time.

In our 5-step approach to logical formalization, the translation of the informal axioms into formal logic is probably the least important step. It is also the least complicated step, given that the rational reconstruction of the theory is carried out well. In fact, the most important role of logical formalization, in our view, is to set a quality standard for the rational reconstruction. The logical formalization is the "proof of the pudding"; if it is hard to do, then the rational reconstruction should be reconsidered.

Aim 2. Applying formalization to social science theories

We applied our 5-step formalization approach to an actual social science theory: resource partitioning. Moreover, we analyzed the theory's key concept, competition, in a rigorous fashion. In our opinion, resource partitioning theory benefited from the formalization effort in a number of ways. (1) Six additional assertions were proved to be redundant, which increased the theory's parsimony. The theory resulting from the formalization effort, is "lean" in the sense that it contains precisely the assumptions necessary to derive the theorems. (2) The formalization increased the theory's parsimony by reducing the number of concepts in the theory. Crowding and competition, for example, turned out to have the same meaning, as did arena and resource base. (3) By mutually relating a number of concepts of the theory, its parsimony was increased once more. (4) The formalization forced us to identify a number of background assumptions and meaning postulates, making resource partitioning theory logically sound. (5) By adding an—unrestrictive—assumption to the theory, a new theorem could be derived, explaining a previously unexplained social phenomenon.

(6) The formalization of the concept of competition by means of network-analysis in Chapter 4, resulted in independent operational definitions of competition at the micro level and competition at the macro level. (7) The use of the operational definitions enabled us to actually measure competitive processes. (8) The operational definitions also enabled us to empirically test all steps in the argument of resource partitioning, thereby testing the entire theory, rather than just its implications. (9) In Chapter 5, the logical formalization of Chapter 3, and the conceptual analysis of Chapter 4, were used to formulate a theory of resource partitioning that applies to emerging industries, as well as to mature industries. (10) The formalization pointed out the conceptual differences between resource partitioning theory and density dependence theory. (11) Deleting these differences, by replacing density dependence's concept of competition, enabled us to

connect density dependence theory and resource partitioning theory.

Our formalization of the concept of competition has, aside from the theoretical implications sketched above, also implications for (future) empirical research. The operational definitions that resulted from the network-representation apply perfectly to a huge database of competitive relations that everyone has at his disposal: the Internet. We succeeded in collecting data about the competitive relations in 9316 pairs of firms, constituting an entire industry, at 28 assessments. We did so by writing a simple computer program that asked the Internet—in particular, search engine Northernlight—for the data we needed. This "easy" way of collecting data may be duplicated for other industries, and other competitive domains; in particular those with a significant Internet exposure, such as software manufacturers, Internet retailers, car types, rock bands, writers, political parties, candidates running for office, universities, and more.

Aim 3. Elucidating social processes

In the Internet search engine market, specialists proliferated during the emergence of the industry. As far as we know, this phenomenon has not been observed in other industries. By expanding the domain of resource partitioning theory to include emerging industries, we were able to attribute a part of the specialist proliferation to resource partitioning processes. A significant part of the unexpectedly high specialist inflow, however, we could not explain. We have looked at resource space dimensionality as a possible explanation, but failed to find any supporting evidence. Nevertheless, we should not drop this way of explaining specialist emergence altogether. The curse of dimensionality implies that eventually the demand for "main stream" products decreases if consumer tastes become more sophisticated, and consumers start taking many different product dimensions into account. The reason is that most consumers in a multidimensional market differ from the main stream on at least one dimension. In such sophisticated markets, resources may be too scattered for generalists to make a living, and (small) specialist firms will become the dominant form.

Such a process certainly seemed to be taking place in the Internet search engine market in our focal period: large, generalist search engines lost ground to small search engines specializing in a particular language area or to a particular geographical domain. Near the end of our focal period, generalist search engine Google entered the market, and, as we know now, had the technology to beat many specialist search engines on their own turf. However, the Internet is still growing too fast for any single search engine to keep up with. Therefore, specialist search engines will remain a viable organizational form, although they have to keep up their technology level, and may be forced to move further out to the market periphery.

Conclusion

Our 5-step approach to (logical) formalization of social science theories helps to produce formal, sound and consistent theories. Moreover, it facilitates the scientific debate about the theories. By applying the approach to resource partitioning theory, the theory's rigor and precision were increased, as well as its parsimony. Moreover, implicit background assumptions were identified and a valuable new theorem was derived. The formalization of the concept of competition by means of a network-representation increased our understanding of competitive processes. Furthermore, it facilitated establishing a link between competitive processes at the micro level of individual firms, to competitive processes at the macro level of industries. By means of the micro-macro link, resource partitioning theory could be connected to density dependence theory. The network-representation also helped to operationally define competition, and to measure competitive processes in the Internet search engine market. During the emergence of this market, the specialist subpopulation started to proliferate. This proliferation can partly be attributed to processes of resource partitioning: specialist search engines moved toward the periphery of the market, and, consequently, improved their life chances.

Bibliography

Andréka, H., J. X. Madarász, I. Németi, G. Sági, and I. Sain (1998). Analyzing the logical structure of relativity theory via model theoretic logic. Manuscript, Mathematical Institute of the Hugarian Academy of Sciences.

Axelrod, R. (1997). *The Complexity of Cooperation*. Princeton, N.J. Princeton University Press.

Ayer, A. (1959). *Logical Positivism*. New York. Free Press.

Balzer, W., C. U. Moulines, and J. D. Sneed (1987). *An Architectonic for Science*. Dordrecht. Reidel.

Barabási, A.-L. and R. Albert (1999). Emerging of scaling in random networks. *Science 286*, 509–512.

Barnett, W. (1997). The dynamics of competitive intensity. *Administrative Science Quarterly 42*, 128–160.

Barnett, W. and G. Carroll (1987). Competition and mutualism among early telephone companies. *Administrative Science Quarterly 30*, 400–421.

Barnett, W. P. and T. L. Amburgey (1990). Do larger organizations generate stronger competition? In J. Singh (Ed.), *Organizational Evolution*. Newbury Park, CA. Sage.

Barney, J. B. and W. G. Ouchi (1986). *Organizational Economics*. San Francisco. Jossey-Bass.

Barron, D. N. (1999). The structuring of organizational populations. *American Sociological Review 64*, 421–445.

Barwise, J. and J. Etchemendy (1999). *Language, Proof and Logic*. Stanford. CSLI.

Baum, J. A. (1995). Handbook of organizational studies. In S. R. Clegg, C. Hardy, and W. North (Eds.), *Organizational Ecology*. Sage.

Baum, J. A. C. and H. A. Haveman (1997). Love thy neighbor? differentiation and agglomeration in the manhattan hotel industry, 1898–1990. *Administrative Science Quarterly 42*, 304–338.

Baum, J. A. C. and S. J. Mezias (1992). Localized competition and organizational failure in the manhattan hotel industry, 1898–1990. *Administrative Science Quarterly 37*, 580–604.

Beth, E. W. (1962). *Formal Methods*. Dordrecht. Reidel.

Blalock, H. M. (1969). *Theory Construction*. Englewood Cliffs. Prentice-Hall.

Boone, C., V. Bröcheler, and G. R. Carroll (2000). Custom service: Application and test of resource partitioning theory among Dutch auditing firms from 1896 to 1992. *Organization Studies 21*(2), 355–381.

Boone, C., A. van Witteloostuijn, and G. R. Carroll (2002). Resource distributions and market partitioning: Dutch daily newspapers, 1968 to 1994. *American Sociological Review 67*, 408–431.

Borenstein, S. and J. Netz (1999). Why do all the flights leave at 8 am?: Competition and departure-time differentiation in airline markets. *International Journal of the Industrial Organization 17*, 611–640.

Borgatti, S. P. and M. G. Everett (1999). Models of core/periphery structures. *Social Networks 21*, 375–395.

Brüderl, J., P. Preisendörfer, and R. Ziegler (1992). Survival chances of newly founded business organizations. *American Sociological Review 57*, 227–242.

Bruggeman, J. (1997a). *Formalizing Organizational Ecology*. Ph. D. thesis, Amsterdam, ILLC DS-1996-4.

Bruggeman, J. (1997b). Niche width theory reappraised. *Journal of Mathematical Sociology 22*, 201–220.

Bruggeman, J., G. Carnabuci, and I. E. Vermeulen (2002). A note on structural holes theory and niche overlap. *Social Networks*. refereed.

Bruggeman, J. and I. E. Vermeulen (2002). A logical toolkit for theory (re)construction. *Sociological Methodology 32*, 183–217.

Carnap, R. (1928). *Der logische Aufbau der Welt*. Berlin. Weltkreis.

Carroll, G. R. (1985). Concentration and specialization: Dynamics of niche width in populations of organizations. *American Journal of Sociology 90*, 1262–1283.

Carroll, G. R. and M. Hannan (1989a). Density delay in the evolution of organizational populations: A model and five empirical tests. *Administrive Science Quarterly 34*, 411–430.

Carroll, G. R. and M. Hannan (1989b). Density dependence in the evolution of newspaper populations. *American Sociologal Review 54*, 524–541.

Carroll, G. R. and M. T. Hannan (Eds.) (1995). *Organizations in Industry.* Oxford. Oxford University Press.

Carroll, G. R. and M. T. Hannan (2000). *The Demography of Corporations and Industries.* Princeton. Princeton University Press.

Carroll, G. R. and A. Swaminathan (1991). Density dependent organizational evolution in the american brewing industry from 1633 to 1988. *Acta Sociologica 34*(3), 155–175.

Carroll, G. R. and A. Swaminathan (2000). Why the microbrewery movement? Organizational dynamics of resource partitioning in the U.S. brewing industry. *American Journal of Sociology 106*(3), 715–762.

Chang, C. and H. J. Keisler (1990). *Model Theory* (3rd ed.). Amsterdam. North-Holland.

Coleman, J. S. (1964). *An Introduction to Mathematical Sociology.* New York. Free Press.

Coleman, J. S. (1990). *Foundations of Social Theory.* Cambridge, MA. Harvard University Press.

Cox, T. F. and M. A. A. Cox (1994). *Multidimensional Scaling.* Chapman & Hall, London UK.

De Groot, A. D. (1961). *Methodologie: Grondslagen van onderzoek en denken in de gedragswetenschappen.* 's-Gravenhage. Mouton.

Debreu, G. (1959). *Theory of Value.* New Haven. Yale University Press.

Eaton, C. and R. Lipsey (1975). The principle of minimum differentiation reconsidered: Some new developments in the theory of spatial competition. *Review of Economic Studies 42*, 27–49.

Enderton, H. B. (1972). *A Mathematical Introduction to Logic.* New York. Academic Press.

Enderton, H. B. (1977). *Elements of Set Theory.* New York. Academic Press.

Farquharson, R. (1969). *Theory of Voting.* Oxford. Blackwell.

Feigenbaum, E. A. and J. Feldman (Eds.) (1963). *Computers and Thought.* New York. McGraw-Hill.

Fisher, A. (1988). *The Logic of Real Arguments.* Cambridge. Cambridge University Press.

Fitting, M. (1996). *First-Order Logic and Automated Theorem Proving* (2nd ed.). New York. Springer.

Freeman, J. and M. T. Hannan (1983). Niche width and the dynamics of organizational populations. *American Journal of Sociology 88*, 1116–1145.

Freeman, J. and A. Lomi (1994). Resource partitioning and foundings of banking cooperatives in Italy. In J. A. Baum and J. V. Singh (Eds.), *Evolutionary Dynamics of Organizations*, pp. 269–293. New York. Oxford U.P.

Freese, L. (1980). Formal theorizing. *Annual Review of Sociology 6*(187-212).

Frege, G. (1961). *Funktion, Begriff, Bedeutung*. Göttingen. Vandenhoeck.

Gabbay, D. and F. Guenthner (Eds.) (1983-1989). *Handbook of Philosophical Logic*. Dordrecht. Reidel.

Gamut, L. (1991). *Logic, Language and Meaning*. Chicago. Chicago U.P.

Gandal, N. (2001). The dynamics of competition in the internet search engine market. *International Journal of Industrial Organization 19*(1103-1117).

Gardner, M. (1983). *Logic Machines and Diagrams*. Brighton. Harvester Press.

Gershenfeld, N. (1999). *The Nature of Mathematical Modeling*. Campbridge University Press.

Gilbert, N. and K. G. Troitzsch (1999). *Simulation for the Social Scientist*. Open University Press.

Groenboom, R., E. Saaman, E. Rotterdam, and G. R. de Lavalette (1996). Formalizing anaesthesia: A case study in formal specification. In M. Gaudel and J. Woodcock (Eds.), *Industrial Benefit and Advances in Formal Methods*, Number 1051 in Lecture Notes in Computer Science, pp. 120–139. New York. Springer.

Hage, J. (1965). An axiomatic theory of organizations. *Administrative Science Quarterly 10*(3), 289–320.

Hage, J. (Ed.) (1994). *Formal Theory in Sociology*. New York. State U.P.

Halmos, P. R. (1960). *Naive Set Theory*. New York. Springer.

Hannan, M., J. Ranger-Moore, and J. Banaszak-Holl (1990). Competition and the evolution of organizational size distributions. In J. V. Singh (Ed.), *Organizational Evolution: New Directions*, pp. 246–268. Newbury Park. Sage.

Hannan, M. T. (1997). On logical formalization of theories from organizational ecology. In A. E. Raftery (Ed.), *Sociological Methodology*, pp. 145–149. Oxford. Blackwell.

Hannan, M. T. (1998). Rethinking age dependence in organizational mortality: Logical formalizations. *American Journal of Sociology 104*, 126–164.

Hannan, M. T. and G. R. Carroll (1992). *Dynamics of Organizational Populations*. New York. Oxford University Press.

Hannan, M. T., G. R. Carroll, E. A. Dundon, and J. C. Torres (1995). Organizational evolution in a multinational context: Entries of automobile manufacturers in Belgium, Britain, France, Germany, and Italy. *American Sociological Review 60*, 509–528.

Hannan, M. T. and J. Freeman (1977). The population ecology of organizations. *American Journal of Sociology 82*, 929–964.

Hannan, M. T. and J. Freeman (1984). Structural inertia and organizational change. *American Sociological Review 49*, 149–164.

Hannan, M. T. and J. Freeman (1989). *Organizational Ecology*. Cambridge, MA. Harvard University Press.

Hansson, S. O. (2000). Formalization in philosophy. *The Bulletin Of Symbolic Logic 6*, 162–175.

Hawking, D. and N. Craswell (2001). Measuring search engine quality. *Journal of Information Retrieval 4*, 33–59.

Hempel, C. G. (1966). *Philosophy of Natural Science*. Englewood Cliffs. Prentice-Hall.

Hodges, W. (1983). Elementary predicate logic. In D. Gabbay and F. Guenthner (Eds.), *Handbook of Philosophical Logic*, Volume 1. Dordrecht. Reidel.

Hodges, W. (1998). An editor recalls some hopeless papers. *The Bulletin of Symbolic Logic 4*, 1–16.

Homans, G. C. (1967). *The Nature of Social Science*. New York. Harcourt, Brace and World.

Homans, G. C. (1980). Discovery and the discovered in social theory. In H. Blalock (Ed.), *Sociological Theory and Research*, pp. 17–22. New York. Free Press.

Hotelling, H. (1929). Stability in competition. *The Economic Journal 39*, 41–57.

Ijiri, Y. and H. A. Simon (1977). *Skew Distributions and the Sizes of Business Firms*. Amsterdam. North-Holland.

Jovanovic, B. (2001). Fitness and age: Review of carroll and hannan's demography of corporations and industries. *Journal of Economic Literature 39*, 105–119.

Kamps, J. (1998). Formal theory building using automated reasoning tools. In A. G. Cohn, L. K. Schubert, and S. C. Shapiro (Eds.), *Principles of Knowledge Representation and Reasoning: Proceedings of the Sixth International Conference (KR'98)*, pp. 478–487. Morgan Kaufmann Publishers, San Francisco CA.

Kamps, J. (1999a). On criteria for formal theory building: Applying logic and automated reasoning tools to the social sciences. In J. Hendler and D. Subramanian (Eds.), *Proceedings of the Sixteenth National Conference on Artificial Intelligence (AAAI-99)*, pp. 285–290. AAAI Press/The MIT Press, Menlo Park CA.

Kamps, J. (1999b). On the process of axiomatizing scientific theories: Using justification criteria in the context of discovery. In *Proceedings of the AISB'99 Symposium on Scientific Creativity*, pp. 49–58. The Society for the Study of Artificial Intelligence and Simulation of Behaviour.

Kamps, J. and M. Masuch (1997). Partial deductive closure: Logical simulation and management science. *Management Science 43*(9), 1229–1245.

Kamps, J. and L. Pólos (1999). Reducing uncertainty: A formal theory of *organizations in action*. *American Journal of Sociology 104*, 1776–1812.

Kiser, E. (1997). Comment: Evaluating qualitative methodologies. In A. E. Raftery (Ed.), *Sociological Methodology*, pp. 151–158. Oxford. Blackwell.

Knoke, D. (2000). Book review: The demography of corporations and industry. *American Journal of Sociology 106*, 838–840.

Kyburg, J. and H. E. (1968). *Philosophy of Science: A Formal Approach*. New York. Macmillan.

Lakatos, I. (1976). *Proofs and Refutations*. Cambridge. Cambridge U.P.

Lawrence, S. and C. L. Giles (1998). Searching the world wide web. *Science 280*, 98–100.

Lawrence, S. and C. L. Giles (1999). Accessibility of information on the web. *Nature 400*, 107–109.

Levins, R. (1968). *Evolution in Changing Environments*. Princeton, N.J. Princeton University Press.

Lomi, A. (1995). The population and community ecology of organizational founding: Italian cooperative banks. *European Sociological Review 11*, 75–98.

Maurer, S. M. and B. A. Huberman (2000). Competitive dynamics of web sites. Technical report, Xerox Palo Alto Research Center.

McPherson, M. (1983). An ecology of affiliation. *American Sociological Review 48*, 519–532.

Mezias, J. M. and S. J. Mezias (2000). Resource partitioning, the founding of specialist firms, and innovation: The american feature film industry, 1912-1929. *Organization Science 11*(3), 306–322.

Mitchell, W. (1995). Medical diagnostic imaging manufacturers. In G. R. Carroll and M. T. Hannan (Eds.), *Organizations in Industry*. Oxford. Oxford University Press.

Neurath, O. (1970). Foundations of the social sciences. In O. Neurath, R. Carnap, and C. Morris (Eds.), *Foundations of the Unity of Science*, Volume 2, pp. 1–51. Chicago. Chicago University Press.

Peirce, C. S. (1931). *Collected Papers*. Cambridge. Harvard University Press.

Péli, G. (1993). A logical formalization of population dynamics: The density dependence model of organizational ecology. Technical report, CCSOM, University of Amsterdam.

Péli, G. (1997). The niche hiker's guide to population ecology: A logical reconstruction of organizational ecology's niche theory. In A. E. Raftery (Ed.), *Sociological Methodology*, pp. 1–46. Oxford. Blackwell.

Péli, G., J. Bruggeman, M. Masuch, and B. Ó Nualláin (1994). A logical approach to formalizing organizational ecology. *American Sociological Review 59*, 571–593.

Péli, G. and M. Masuch (1997). The logic of propagation strategies: Axiomatizing a fragment of organizational ecology in first-order logic. *Organization Science 8*, 310–331.

Péli, G. and B. Nooteboom (1999). Market partitioning and the geometry of the resource space. *American Journal of Sociology 104*, 1132–1153.

Plato (1987). *Theaetetus*. Harmondsworth. Penguin.

Podolny, J. M., T. E. Suart, and M. T. Hannan (1996). Networks, knowledge, and niches: Competition in the worldwide semiconductor industry, 1984-1991. *American Journal of Sociology 102*, 659–689.

Pólya, G. (1945). *How to solve it*. Princeton. Princeton U.P.

Popper, K. (1959). *The logic of scientific discovery*. London. Hutchinson.

Quine, W. V. O. (1961). *From a Logical Point of View*. New York. Harper and Row.

Quine, W. V. O. (1986). *Philosophy of Logic* (2nd ed.). Cambridge, MA. Harvard University Press.

Quine, W. V. O. (1987 (1990)). *Quiddities*. Harmondsworth. Penguin.

Rapoport, A. (1959). Uses and limitations of mathematical models in social sciences. In L. Gross (Ed.), *Symposium on Sociological Theory*. Evanston, IL. Row Peterson.

Sastry, M. A. (1997). Problems and paradoxes in a model of punctuated organizational change. *Administrative Science Quarterly 42*, 237–275.

Schelling, T. C. (1978). *Micromotives and Macrobehavior* (2 nd ed.). New York. Norton.

Schmalensee, R. and R. Willig (Eds.) (1989). *The Handbook of Industrial Organization*. Elsevier Science, The netherlands.

Shapiro, C. and H. R. Varian (1999). *Information Rules*. Harvard Bussiness School Press.

Shepherd, W. (1987). Concentration. In J. E. *et al.* (Ed.), *The New Palgrave*, pp. 563–564. London. Macmillan.

Simon, H. A. (1954). Some strategic considerations in the construction of social science models. In P. F. Lazarsfeld (Ed.), *Mathematical Thinking in the Social Sciences*, pp. 388–415. The Free Press, Glencoe IL.

Sohn, M.-W. (2001). Distance and cosine measures of niche overlap. *Social Networks 23*, 141–165.

Sonnenreich, W. and T. Macinta (1998). *Web Developer.com Guide to Search Engines*. Wiley Computer Publishing, New York.

Spencer Brown, G. (1957). *Probability and Scientific Inference*. London. Longmans.

Steyvers, M. (2002). Multidimensional scaling. In L. Nadel (Ed.), *Encyclopedia of Cognitive Science*. Macmillan Reference Ltd.

Stinchcombe, Arthur, L. (1968). *Constructing Social Theories*. New York.. Harcourt.

Suppes, P. (1968). The desirability of formalization in science. *Journal of Philosophy 65*, 651–664.

Sutton, R. I. and B. M. Staw (1995). What theory is *not*. *Administrative Science Quarterly 40*, 371–384.

Swaminathan, A. (1995). The proliferation of specialist organizations in the American wine industry, 1941-1990. *Administrative Science Quarterly 40*, 653–680.

Swaminathan, A. and G. R. Carroll (1995). Beer brewers. In G. R. Carroll and M. T. Hannan (Eds.), *Organizations in Industry*. Oxford. Oxford University Press.

Tarski, A. (1941). *Introduction to Logic*. New York. Oxford University Press.

Thornton, P. H. (1999). Concentrated markets in cultural industries: The decline of new firms in the higher education publishing industry. Paper presented at the American Sociological Association.

Van Benthem, J. (1982). The logical study of science. *Synthese 51*, 431–472.

Van Benthem, J. (1994). Logic and argumentation. Invited lecture at the Third Conference of the International Society for the Study of Argumentation, Amsterdam.

Van Benthem, J. and A. Ter Meulen (Eds.) (1997). *Handbook of Logic and Language*. North-Holland. Amsterdam.

Van Dalen, D. (1994). *Logic and Structure* (3rd ed.). Berlin. Springer-Verlag.

Vermeulen, I. E. (2002a). An extension to the network-representation of competition in industries. Draft. Available at: `http://www.science.uva.nl/~ivar/pub/comp_indu_ext.pdf` .

Vermeulen, I. E. (2002b). In search of a niche: Competitive dynamics in the market of internet search engines. Paper presented at the International Sunbelt Social Network Conference XXII.

Vermeulen, I. E. and J. Bruggeman (2001). The logic of organizational markets: Thinking through resource partitioning theory. *Computational and Mathematical Organization Theory 7*, 87–111.

Wade, J. (1996). A community-level analysis of sources and rates of technological variation in the microprocessor market. *Academy of Management Journal 39*, 1218–1244.

Wasserman, S. and K. Faust (1994). *Social Network Analysis*. Cambridge University Press.

Wegman, E. J. (1990). Hyperdimensional data analysis using parallel coordinates. *Journal of the American Statistical Association 85*, 664–675.

Woodger, J. (1937). *The Axiomatic Method in Biology*. Cambridge. Cambridge University Press.

Wos, L. (1996). *The Automation of Reasoning*. San Diego. Academic Press.

Wos, L., R. Overbeek, E. Lusk, and J. Boyle (1991). *Automated Reasoning* (2nd ed.). New York. McGraw-Hill.

Young, R. (1988). Is population ecology a useful paradigm for the study of organizations? *American Journal of Sociology 94*, 1–24.

Samenvatting

Sociale wetenschappers presenteren hun ideeën en theorieën doorgaans in "natuurlijke taal," bijvoorbeeld in Engelse of Nederlandse teksten. Deze manier van presenteren heeft als voordeel dat het voor bijna iedereen mogelijk is om van de gepresenteerde ideeën en theorieën kennis te nemen en ze te begrijpen. Het nadeel is dat natuurlijke taal notoir ambigu is, zowel in conceptuele als in logische zin.

In een aantal recente studies zijn enkele in natuurlijke taal gepresenteerde sociaal-wetenschappelijke theorieën met behulp van formele logica onderzocht. De nadruk lag hierbij op theorieën uit de organisatie ecologie. Organisatie ecologie (OE) houdt zich bezig met de dynamiek in populaties van organisaties. Een belangrijke aanname van OE is dat organisaties slechts in beperkte mate beschikken over informatie over hun omgeving, en dat ze—mede als gevolg daarvan—ook slechts in beperkte mate rationeel handelen. Succes en ondergang van organisaties wordt volgens OE dan ook veeleer verklaard door toeval en omgevingsfactoren, dan door de kwaliteit van het ondernemersschap.

OE is herhaaldelijk gekozen als domein van formeel-logisch onderzoek omdat het een in hoge mate ontwikkelde, samenhangende, en relatief formele verzameling van theorieën betreft. Desondanks zijn met behulp van formele logica in een aantal theorieën redeneerfouten, ontbrekende achtergrondaannames en inconsistenties blootgelegd. Bovendien zijn met behulp van logica redundante aannames aangewezen, en werd de onderlinge samenhang van de theorieën verbeterd. Hoewel formeel-logisch onderzoek in de sociale wetenschappen vooralsnog dus een succesvolle benadering is gebleken, is er tot nu toe weinig aandacht geweest voor de methodologie ervan.

Hoofdstuk 1. Dit hoofdstuk vat de doelen van dit proefschrift als volgt samen:

1. Het ontwikkelen van een systematische benadering van het logisch formaliseren van sociaal-wetenschappelijke theorieën.

2. Het toepassen van logische formalisering op sociaal-wetenschappelijke theorieën, in het bijzonder op de organisatie-ecologische deeltheorie *resource*

partitioning.

3. Het middels formalisering verbeteren van verklaringen van sociale processen, in het bijzonder van de concurrentiedynamiek in de markt van internetzoekmachines.

Hoofdstuk 2. In het tweede hoofdstuk van dit proefschrift wordt een systematische vijf-stappenbenadering van het (logisch) formaliseren van sociaal-wetenschappelijke theorieën gepresenteerd.

De eerste stap in deze benadering is het identificeren van de "kerntheorie." Deze bestaat uit een aantal zinnen: de hoofdconclusie(s) van een theorie en de premissen die nodig zijn om tot de conclusie(s) te komen. Vervolgens wordt, in de tweede stap, de betekenis van de concepten in de kerntheorie vastgelegd. Bovendien worden eventuele overeenkomsten of relaties tussen de concepten geanalyseerd en vastgelegd. In de derde stap wordt de logische structuur van de zinnen die deel uitmaken van de kerntheorie onderzocht. Het resultaat van deze stap is een verzameling informele axioma's: zinnen met een duidelijke logische structuur, en met zo weinig mogelijk concepten. Alle axioma's krijgen bovendien een rol toebedeeld in de redenering: premisse of conclusie.

De eerste drie stappen van de vijf-stappenbenadering van het logisch formaliseren vormen tezamen de "rationele reconstructie." Als deze reconstructie goed is uitgevoerd, is de vierde stap, de daadwerkelijke formalisering in een logische taal, betrekkelijk eenvoudig. Het resultaat van de vierde stap is een verzameling formele axioma's. In de vijfde stap, ten slotte, wordt getracht om, eventueel met behulp van een geautomatiseerde *model generator* en *theorem prover*, de consistentie en geldigheid van de formele theorie vast te stellen. Als dat niet lukt, worden de voorafgaande stappen van het formaliseringsprocess opnieuw doorlopen; formaliseren is een iteratief proces.

Hoofdstuk 3. In dit hoofdstuk wordt de vijf-stappenbenadering toegepast op resource partitioning—een deeltheorie van OE. Deze theorie stelt, in tegenstelling tot de klassieke economische theorie, dat marktconcentratie gunstig is voor kleine gespecialiseerde organisaties en ongunstig voor (middel-)grote generalisten. De redenering is dat marktconcentratie enerzijds zorgt voor het verdwijnen van middelgrote organisaties en anderzijds grote organisaties naar het centrum van de markt drijft. Als gevolg hiervan ontstaan er niches in de periferie van de markt, waar kleine gespecialiseerde organisaties zich kunnen vestigen. De theorie is onderbouwd met empirisch onderzoek in verschillende markten.

De logische formalisering van resource partitioning theorie brengt aan het licht, dat de stelling dat markconcentratie de levenskansen van (middel-)grote generalisten en kleine specialisten respectievelijk negatief en positief beïnvloedt, niet door de argumentatie wordt onderbouwd. Beide effecten worden in de theorie teruggeleid tot het bestaan van schaalvoordelen. Schaalvoordelen kunnen weliswaar markconcentratie veroorzaken, maar zijn daar zeker niet gelijk aan.

Bovendien blijkt een aantal aannames van resource partitioning theorie overbodig te zijn; hetzelfde geldt voor zes erkende additionele verklaringen voor resource partitioning, aangedragen in gerelateerde literatuur. Een aantal impliciete achtergrondaannames moest aan de theorie worden toegevoegd om de argumentatie sluitend te maken. Ten slotte blijkt dat door het toevoegen van de relatief zwakke aanname dat (middel-)grote generalisten meer concurrentiedruk te verwerken hebben dan kleine specialisten, de empirische bevinding verklaard kan worden dat het aantal de markt betredende specialisten doorgaans groter is dan het aantal de markt verlatende generalisten.

Behalve aan het vergroten van de kennis over het mechanisme achter resource partitioning, draagt de logische formalisatie bij aan het "afslanken" van resource partitioning theorie door het aantal concepten in de theorie te verkleinen. Bovendien draagt de formalisatie bij aan de operationaliseerbaarheid en meetbaarheid van enkele begrippen uit de theorie, en aan het relateren van de theorie aan andere theorieën over concurrentie.

Hoofdstuk 4. Het sleutelbegrip van resource partitioning theorie is concurrentie. In Hoofdstuk 4 wordt dit begrip conceptueel onder de loep genomen. In feite gaat het hierbij om een uitgebreide conceptuele analyse, zoals in stap 2 van de in Hoofdstuk 2 gepresenteerde formaliseringsbenadering wordt toegepast.

Het resultaat van de conceptuele analyse is een netwerk-representatie van concurrentie; concurrentie wordt beschouwd als een netwerk, bestaande uit organisaties en hun onderlinge concurrentierelaties.

De netwerkbenadering van concurrentie gaat ervan uit dat als twee organisaties aanspraak maken op een *resource*—een hulpbron, bijvoorbeeld een klant, of een vergunning—zij een concurrentierelatie hebben. De intensiteit van deze relatie wordt bepaald door het aantal resources waar door beide organisaties om geconcurreerd wordt. 'Concurrentiedruk' van organisatie I op J wordt gedefinieerd als het gedeelte van J's resources waar I ook aanspraak op maakt. De intensiteit van een concurrentierelatie en de daarmee gepaard gaande concurrentiedruk zijn gerelateerd aan het organisatie-ecologische kernbegrip *niche overlap*. Door de concurrentiedruk van alle organisaties op organisatie J op te tellen, kan de totale concurrentiedruk op J worden berekend. Door de concurrentiedruk op alle organisaties in een markt op te tellen, wordt de totale concurrentie in een markt berekend.

De netwerkrepresentatie verbindt concurrentie op het micro-niveau (dat van individuele actoren) met concurrentie op het macro-niveau—dat van de markt; door het meten van concurrentierelaties tussen individuele actoren kunnen zo uitspraken worden gedaan over concurrentiedynamiek in een markt. De verbinding van het micro- en het macro-niveau van concurrentie zorgt er ook voor dat theorieën over de verschillende niveaus van concurrentie aan elkaar kunnen worden gekoppeld.

Uit de netwerkrepresentatie van concurrentie kan worden afgeleid, dat de grootte van een organisatie positief gerelateerd is aan de mate van concurrentiedruk die een organisatie uitoefent. De grootte van organisaties heeft echter nauwelijks invloed op de concurrentiedruk die organisaties te verduren krijgen; die is met name gerelateerd aan de positionering ten opzichte van andere organisaties.

Om de bruikbaarheid van de netwerkbenadering van concurrentie aan te tonen, wordt de intensiteit van de paarsgewijze concurrentierelaties in een populatie van 137 internet-zoekmachines gemeten, per kwartaal, over een periode van zeven jaar. Een aantal hypotheses over de te verwachten concurrentiedynamiek in opkomende industrieën worden aan de hand van de verzamelde data getest. De voornaamste bevinding is, dat in de populatie van internet-zoekmachines concurrentie eerst in toenemende mate toeneemt, maar dat vervolgens, door het betreden van de markt van een grote groep kleine specialisten, de toename van marktconcurrentie stagneert. Dit resulteert in een S-vormige toename van de concurrentie in de zoekmachinemarkt. Verder wordt het theoretische resultaat bevestigd dat grote organisaties meer concurrentiedruk uitoefenen dan kleine organisaties, en ongeveer dezelfde concurrentiedruk te verduren krijgen.

Hoofdstuk 5. Dit hoofdstuk beschrijft het empirisch onderzoek van de markt van internet-zoekmachines, op basis van de theoretische bevindingen uit de eerdere hoofdstukken van het proefschrift. Een markant onderdeel van de opkomst van de zoekmachinemarkt is de relatieve groei van het aantal kleine specialisten in de markt. Vanaf 1997 is meer dan 90% van de zoekmachines die de markt betreedt een specialist. Om de opkomst van de specialisten in de groeiende populatie te verklaren, wordt geprobeerd twee organisatie-ecologische theorieën, density dependence theorie, die zich bezig houdt met de opkomst van populaties van organisaties, en de eerder genoemde resource partitioning theorie, aan elkaar te verbinden. Beide theorieën baseren zich op verschillende aannames en hebben betrekking op verschillende 'levensfasen' van de populatie. Aangetoond wordt dat door de vervanging van het concurrentiebegrip in beide theorieën door de in Hoofdstuk 4 ontwikkelde netwerkrepresentatie van concurrentie, de theorieën aan elkaar kunnen worden verbonden.

De analyse van de empirische data toont aan dat legitimatie in de zoekmachinemarkt in toenemende mate toeneemt, hetgeen niet overeenkomt met wat door density dependence theorie wordt aangenomen. De concurrentie in de zoekmachinemarkt doorloopt een, ook al afwijkend, S-vormig patroon. Er kan geen relatie worden vastgesteld tussen marktconcurrentie, legitimatie en de in- en uitstroom van organisaties, zoals in density dependence theorie wordt verondersteld. Kennelijk kan density dependence theorie de waargenomen dynamiek in de zoekmachinemarkt niet verklaren.

Resource partitioning theorie daarentegen blijkt wel een verklarende waarde te hebben. Overeenkomstig met de aannames in de theorie, neemt de groei van het aantal generalisten in de zoekmachinemarkt af, terwijl de concurrentie onder generalisten toeneemt; dit laatste fenomeen lijkt op wat resource partitioning theorie de ‘trek naar het centrum van de markt’ noemt. Als gevolg van deze beweging van generalisten, maar vooral doordat specialisten van het centrum van de markt af bewegen, neemt de concurrentie tussen generalisten en specialisten af. Daardoor neemt de concurrentiedruk op specialisten af, wat op zijn beurt weer resulteert in een toenemende de groei van het aantal specialisten, zoals waargenomen in 1999 en 2000.

Hiermee is de relatieve groei van het aantal specialisten in de zoekmachinemarkt tussen 1997 en 1999 echter nog niet verklaard; daarom wordt nog gekeken naar een andere verklaring uit de literatuur, het *sphere packing* model, dat een toenemende dimensionaliteit van de z.g. *resource space* als mogelijke oorzaak van het succes van kleine specialisten ziet. De oorspronkelijke verklaring blijkt geen betrekking te hebben op het onderzochte empirische domein. Na de verklaring voor het onderzoeksdomein te hebben aangepast, blijkt dat van een toenemende dimensionaliteit van de zoekmachinemarkt in de onderzochte periode geen sprake is. De groei van de zoekmachinepopulatie, alsmede de differentiatie van kleine specialisten, vindt plaats in een laagdimensionale markt; de relatieve toename van het aantal specialisten kan dus niet worden teruggeleid tot een hogere dimensionaliteit van de zoekmachinemarkt.

Conclusies. De in dit proefschrift geïntroduceerde vijf-stappenbenadering van het logisch formaliseren van sociaal-wetenschappelijke theorieën heeft geholpen resource partitioning theorie sluitend en consistent te maken. Daarbij vergrootte de formalisering het begrip van de theorie. Bovendien was de geformaliseerde versie van resource partitioning theorie compacter, duidelijker, en preciezer dan het origineel, en kon uit de theorie een belangrijke nieuwe voorspelling worden afgeleid.

De conceptuele analyse van het begrip concurrentie—met behulp van netwerkanalyse-technieken—vergrootte het begrip van concurrentieprocessen. Daarbij werd een directe link tussen concurrentieprocessen op micro- en op macro-niveau bewerkstelligd. Deze link vergemakkelijkte het samenvoegen van resource partitioning theorie met density dependence theorie, alsmede het meten van concurrentieprocessen.

Concurrentieprocessen in de internet-zoekmachinemarkt werden empirisch onderzocht. Dit leidde tot de ontdekking dat de waargenomen opkomst van specialisten in deze groeiende markt ten minste gedeeltelijk kon worden toegeschreven aan resource partitioning: specialisten differentieerden zich door hun heil te zoeken in meer perifere gedeeltes van de markt; op die manier vergrootten ze hun levenskansen.

Titles in the ILLC Dissertation Series:

ILLC DS-1999-01: **Jelle Gerbrandy**
Bisimulations on Planet Kripke

ILLC DS-1999-02: **Khalil Sima'an**
Learning efficient disambiguation

ILLC DS-1999-03: **Jaap Maat**
Philosophical Languages in the Seventeenth Century: Dalgarno, Wilkins, Leibniz

ILLC DS-1999-04: **Barbara Terhal**
Quantum Algorithms and Quantum Entanglement

ILLC DS-2000-01: **Renata Wassermann**
Resource Bounded Belief Revision

ILLC DS-2000-02: **Jaap Kamps**
A Logical Approach to Computational Theory Building (with applications to sociology)

ILLC DS-2000-03: **Marco Vervoort**
Games, Walks and Grammars: Problems I've Worked On

ILLC DS-2000-04: **Paul van Ulsen**
E.W. Beth als logicus

ILLC DS-2000-05: **Carlos Areces**
Logic Engineering. The Case of Description and Hybrid Logics

ILLC DS-2000-06: **Hans van Ditmarsch**
Knowledge Games

ILLC DS-2000-07: **Egbert L.J. Fortuin**
Polysemy or monosemy: Interpretation of the imperative and the dative-infinitive construction in Russian

ILLC DS-2001-01: **Maria Aloni**
Quantification under Conceptual Covers

ILLC DS-2001-02: **Alexander van den Bosch**
Rationality in Discovery - a study of Logic, Cognition, Computation and Neuropharmacology

ILLC DS-2001-03: **Erik de Haas**
Logics For OO Information Systems: a Semantic Study of Object Orientation from a Categorial Substructural Perspective

ILLC DS-2001-04: **Rosalie Iemhoff**
Provability Logic and Admissible Rules

ILLC DS-2001-05: **Eva Hoogland**
Definability and Interpolation: Model-theoretic investigations

ILLC DS-2001-06: **Ronald de Wolf**
Quantum Computing and Communication Complexity

ILLC DS-2001-07: **Katsumi Sasaki**
Logics and Provability

ILLC DS-2001-08: **Allard Tamminga**
Belief Dynamics. (Epistemo)logical Investigations

ILLC DS-2001-09: **Gwen Kerdiles**
Saying It with Pictures: a Logical Landscape of Conceptual Graphs

ILLC DS-2001-10: **Marc Pauly**
Logic for Social Software

ILLC DS-2002-01: **Nikos Massios**
Decision-Theoretic Robotic Surveillance

ILLC DS-2002-02: **Marco Aiello**
Spatial Reasoning: Theory and Practice

ILLC DS-2002-03: **Yuri Engelhardt**
The Language of Graphics

ILLC DS-2002-04: **Willem Klaas van Dam**
On Quantum Computation Theory

ILLC DS-2002-05: **Rosella Gennari**
Mapping Inferences: Constraint Propagation and Diamond Satisfaction

ILLC DS-2002-06: **Ivar Vermeulen**
A Logical Approach to Competition in Industries