STRUCTURAL MARKET CHANGES AND STRATEGIC ADAPTATION ALONG THE VALUE CHAIN: THEORETICAL PERSPECTIVES AND EMPIRICAL EVIDENCE

Inauguraldissertation

zur Erlangung des akademischen Grades eines

Doktors der Wirtschaftswissenschaften durch die

Wirtschaftswissenschaftliche Fakultät

der Westfälischen Wilhelms-Universität Münster

WESTFÄLISCHE
WILHELMS-UNIVERSITÄT
MÜNSTER

vorgelegt von

CARL HENDRIK SCHMALE

aus Ochtrup

Münster, 2010

Erster Berichterstatter:	Prof. Dr. Thomas Ehrmann
Zweiter Berichterstatter:	Prof. Dr. Alexander Dilger
Dekan:	Prof. Dr. Stefan Klein
Tag der mündlichen Prüfung:	06. Juli 2010

ACKNOWLEDGEMENTS

I wish to thank a number of people whose support has been of extraordinary value to me. First and foremost, I wish to thank Prof. Dr. Thomas Ehrmann for supervising my dissertation and for his continuous and enduring support and trust. His invaluable expertise greatly advanced the academic scope of my research. I also extend my thanks to Prof. Dr. Alexander Dilger for his co-supervisory of my thesis and Prof. Dr. Martin Bohl for joining the dissertation committee. Moreover, I am deeply indebted to Prof. Dr. Thomas Ehrmann, Prof. Dr. Alexander Dilger, Prof. Dr. Karl-Hans Hartwig and Dr. Torsten Marner for letting me contribute to our joint projects. I want to thank my friends and colleagues Dr. Brinja Meiseberg, Sandra Stevermüer, Dr. Eugen Scheinker, Dr. Julian Dormann, Dr. Olivier Cochet, Alfred Koch and Jan Piening for inspiring discussions, hours of proof reading and a great time at the Institute of Strategic Management. Also, I wish to thank my family, particularly my parents Claudia and Carl Schmale, and Nina Ringkamp for their support and patience.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

PART A

I. ECONOMIC ROLES, ORGANIZATIONAL STRUCTURE AND VALUE CREATION IN THE FACE OF STRUCTURAL MARKET CHANGE

"Revolutions always come around again. That's why they're called revolutions."
Terry Pratchett, *Night Watch*

Actors in a market assume one or more of the primary economic roles of an industry value chain, i.e. suppliers, producers, distributors and customers, to carry out value-creating activities. Suppliers create product components or provide raw materials, services, or expertise. Producers design and build products or services that aim at meeting customers' specific needs. Distributors enable buyers and sellers to connect, communicate and to transact (Applegate, 2001).

The structuring of economic action, the assignment of roles and the forming of relationships of actors, and hence the positioning of firms in their environment are central subjects of analysis in economic science. In the late 18th century, Adam Smith asserted that the division of labor is limited by the extent of the market. He argued that specialization should become more prevalent with increasing market size, since higher sales justify setting up firms which concentrate on formerly low volume production processes that are subject to increasing returns (Smith, 1776; see Stigler (1951) for a brief discussion and an integration of Smith's conjecture into a theory of the functions of the firm). During the industrial revolution in the late 19th and the early 20th century, the opposite could be witnessed: the advent of large integrated corporations as an answer to the managerial challenges of increasingly complex and high-volume production and distribution processes caused by technological progress, population growth and expanding per capita income (Chandler, 1977).

Reviewing the above illustrated discrepancy of theoretical prediction and fact, Langlois (2003: 352) concludes that it was the imbalance between "the coordination needs of high-throughput technologies and the abilities of contemporary markets and contemporary institutions to meet

those needs" which fostered the rise of large vertically integrated firms, exploiting economies of scale and scope, instead of specialized actors in disintegrated production processes.

These days, markets are again subject to substantial changes which result in adaptations of strategic ends and organizational structure along the value chain (Langlois, 2003). Specifically, economic actors face significant shifts in the competitive landscape that challenge the firms' strategies in the pursuit of obtaining competitive advantage (Hitt, Keats and DeMarie, 1998).

First, developments in public policy change the rules of the game. Intending to realize lower prices for consumers as well as improvements of efficiency and quality, governments take efforts of raising competition by liberalization of formerly regulated industries such as the electricity, air transportation, postal service or railway market. This has significant influence on the economic roles and relationships of actors along the value chain, as de-regulation can imply the disintegration of hitherto vertically integrated production processes, for instance in the case of railways (European Commission, 1996; 2001).

Second, the rapid improvements of information and communication systems, namely, the internet, electronic data interchange (EDI) systems and related network technologies, further lowered communication costs and together with decreasing legal barriers to trade fostered the emergence of integrated markets with high trade volumes (Applegate, 2001; Langlois, 2003). This ubiquitous development, commonly referred to as the "technological revolution" or the rise of the "information age" (Applegate, 2001) greatly impacts the strategic management of firms in the new competitive landscape (Hoskisson, 1999).

Bettis and Hitt (1995) describe the new imperatives that drive value chain configuration and strategic objectives in the new environment. The growing relevance of knowledge accumulation and deployment and decreased transaction costs emerge as main influence factors on strategy formulation. First, the high relevance of knowledge accumulation and deployment follows from the growing technological orientation of industries. That is, firms either make extensive use of technology within the production process or technology is a core element of the business model

(see, e.g. Amit and Zott, 2001). Second, information and communication technologies are lowering transaction costs both within the production process and with customers by eliminating information asymmetries, decreasing asset specificities and by facilitating monitoring which mitigates the effects of opportunistic behavior (Brynjolfsson et al., 1994; Ciborra, 1993; Cordella, 2006; Malone, Yates and Benjamin, 1987).

As a consequence, the general understanding of how to create value, formerly grounded on the assumptions of roles and business processes of the industrial economy, is changing (Amit and Zott, 2001; Applegate, 2001). In particular, two significant trends become apparent: 1) the disintegration of production processes and 2) the re-definition of customers' roles in the value creation process.

As transaction costs are reduced by information and communication technologies, and as knowledge takes up a more prominent role in organizations, externalizing certain functions of the firm has become a favored strategy in many cases. In particular, the above mentioned developments have led scholars and practitioners to see new opportunities for interfirm cooperation (Baum and Oliver, 1991; Gulati, 1995; Pillai, 2006; Powell, Koput and Smith-Doerr, 1996). This insight manifests itself in a growing proliferation of hybrid organizing forms such as strategic alliances, joint ventures, franchising networks and related network forms of organisation (e.g. Parkhe, Wasserman and Ralston, 2006; Podolny and Page, 1998).

Similarly, the boundaries of the firm are fading as customers' economic roles are more frequently re-interpreted. Formerly being essentially the recipient of goods and services, nowadays the customer gets integrated into the value creation process (Prahalad and Ramaswamy, 2004). This integration may take different forms, or degrees. For instance, customers might either serve as valuable information sources by submitting complaints, suggestions or by disclosing preferences through online behavior (Ehrmann, 2003) or they might even be woven into the value creation process by executing tasks which were formerly assigned to the distributor or producer role (e.g. Norman and Ramirez, 1993).

This thesis aims at taking a deeper look at key aspects of the aforementioned trends, i.e. 1) the disintegration of production processes and 2) the re-definition of customers' roles, for the value-creation process and firm performance from a strategic management perspective. Thereby, the chapters intend to critically reflect "common wisdom" about the performance effects of structural adaptations along the value chain. The studies hold implications for the successful design of value creation processes with respect to the assignment of roles, the management of relationships and the choice of goods and services that are exchanged. The basic organizing principle and the specific research questions are outlined in the next section.

II. BASIC ORGANIZING PRINCIPLE

In a way, the value chain can be viewed as a representation of the different economic roles of actors in a market that carry out value creating activities (Figure 1; Applegate, 2001).

Figure 1. *Economic Roles of a Basic Value Chain*

The process of value creation can be analyzed with respect to the dimensions of structure, governance and content of transactions (Amit and Zott, 2001). These dimensions shall serve as a metaphorical framework, intending to facilitate the placement of this thesis' chapters into the broader context of strategic management.[1]

Applied to the context here, *structure* is closely related to the economic roles of a value chain since it describes the transactions that are performed, the actors who perform these transactions and how these actors are linked. Moreover, structure describes the sequencing of transactions as well as the adopted exchange mechanisms that enable transactions. The choice of structure and the assignment of activities to actors can have several crucial implications for value creation. First, actors can hold great potential for value creation (Ehrmann, 2003). For instance, consider the on-line travel portal Expedia.com which features a community-based review platform, thereby letting customers co-create value by providing information on hotels and destinations for other potential buyers. Second, the choice of transaction structure influences the flexibility, adaptability and scalability of transactions. Integrating certain activities implies the set-up of

[1] Originally, Amit and Zott (2001: 511) propose these three dimensions as analytical framework for describing business models. They establish that "A business model depicts the content, structure and governance of transactions designed so as to create value through the exploitation of business opportunities." Although business model analysis in its narrow sense goes beyond the intention of this dissertation, these dimensions shall nevertheless serve as a metaphorical framework accommodating and structuring the specific research foci of this thesis.

specific machinery, the hiring of specialized personnel and so on. So if unforeseen circumstances changed customers' preferences, and demand declined, significant overheads would possibly be left uncovered due to underutilization of capacity. In contrast, if this activity was outsourced, (variable) costs would decline approximately in line with demand (Langlois, 2003). However, structural set-up cuts both ways in terms of scalability and flexibility, as increases in production have to be backed by an increasing supply of intermediate goods. Hence, expansion depends on the capabilities, resources and willingness of outside actors to supply these goods or to perform certain activities.

Transaction *governance* refers to the ways in which transactions, i.e. flows of resources, information and goods, are controlled by the actors (Amit and Zott, 2001). This also refers to the legal forms of organization, contractual design as well as profit allocation mechanisms and hence, the incentive structure of transactions. In this respect, transaction governance is a crucial parameter of economic outcomes as it determines the extent to which potentials for value creation can be translated into actual value. Hence governance mechanisms can either hinder or encourage the timely transfer of information, parties' motivation to act in compliance with prior agreements or to invest specifically, for example.

Content denotes the actual "goods or information that are being exchanged, and to the capabilities and resources that are required to enable the exchange" (Amit and Zott, 2001: 511). The value creation potential of transaction content lies in the combination of complementary and specialized resources and capabilities which differ within an industry (e.g. Barney, 1991; Hamel, 1991; Teece, 1987).

The complex interplay in the choice of transactional structure, governance and content might be illustrated by a very simplified example. Consider the online retailer Amazon.com. In attempting to leverage the technological advantages of online commerce over offline retailing, Amazon had to carefully asses own resources and capabilities as well as the resources and capabilities of the other market participants in assigning activities and shaping the firm specific value chain. On the one hand, Amazon's centralized warehouses and its virtual store allow for an abundant

product variety due to unlimited shelf space and diminishing opportunity costs of listing. However, for this product variety to translate into value, i.e. sales, the heterogeneity of choice has to be made accessible to the customer. Considering only firm internal resources, Amazon would not be able to promote single articles in terms of providing individual advice with respect to the quality of an article or the fit between a customer's preferences and the article's characteristics. Yet, the customer crowd can. In implementing recommender systems which analyze customers' purchase patterns and in providing a platform that allows for user-generated online reviews, Amazon enables value co-creation by customers. By furthermore establishing a system which allows to rate the reviews themselves, Amazon provides incentive structures that promote customers to actually provide review information.

As noted above, the chapters of this dissertation analyze different aspects of structural adaptations along the value chain, thereby considering the general trends of 1) the disintegration of production processes (chapters I and II) and 2) the re-interpretation of the customers' role (chapters III and IV). From a theoretical perspective, each chapter focuses on the structural dimension and – primarily – one of the two other dimensions, i.e. governance and content.[2] The specific research foci and contributions to the literature are outlined below.

[2] As outlined above structural adaptations are – in most cases – connected with both the other dimensions; either by being motivated by resource considerations or because structural adaptations imply adaptations of governance mechanisms. However, in order to provide a concise overview, the chapters are classified according to their primary focus.

III. SUMMARY OF CHAPTERS

1. An Experimental Analysis of Organizational Structure and Specific Investments in Railways*

Chapter I addresses the issue of incentive structures under different organizational designs of the production process in the railway industry. European guidelines suggest a separation of infrastructure provision and transport operation to promote efficiency and to mitigate welfare losses due to potential market foreclosure and discrimination of third-party network access (European Commission, 1991; 2001; Nash and Preston, 1994). However, vertical disintegration can entail substantial disadvantages in terms of value creation because incentive conflicts might lead to severe under-investment in specific assets (e.g. Grossman and Hart, 1986; Hart, 1995; Pittman, 2007; Rothengatter, 2001).

According to modern institutional economics, a vertical separation of infrastructure from operations in network industries is an inferior form of organization if investments are specific and contracts incomplete (Williamson, 1975; Williamson, 1985; Klein, Crawford and Alchian, 1978; Crocker and Masten, 1991). Because specific investments yield significantly lower values or lower gains from trade, when employed in a transaction other than originally intended (Joskow, 2003), the investor bears the risk of being exploited by an opportunistic transaction partner, who will appropriate the difference between the value of the investment in its first and second-best use (quasi-rent) in an ex-post bargaining process. If the investor anticipates the risk of a hold-up and if contractual arrangements to avert hold-ups are hindered by incomplete contracts, he will not undertake the investment at all. In network industries, underinvestment may occur on either side of the transaction.

A large body of *theoretical* research analyzes possible consequences of different organizational designs in railways. Yet, there is only sparse *quantitative* evidence on the effects of separation.

* This chapter originates from joint work with Thomas Ehrmann, Karl-Hans Hartwig and Torsten Marner. For the discussion paper see: Ehrmann, T., Hartwig K.-H., Marner, T. and Schmale, H. (2009), Specific Investments and Ownership Structures in Railways: An Experimental Analysis. Discussion Paper No. 12 of the Institute of Transport Economics, University of Muenster.

This is mainly due to the fact that the settings in the different countries are difficult to compare and that sufficiently large databases necessary for empirical research are only available for few of the relevant variables (Nash and Rivera-Trujillo, 2004). Most previous studies on vertical separation test for economies of scope (Preston, 1996; Growitsch and Wetzel, 2009) and analyze the implications on competition and on efficiency and productivity growth (Nash and Preston, 1994; Bitzan, 2003; Friebel, Ivaldi and Vibes, 2003; Driessen, Lijesen and Mulder, 2006; Sanchez, Monsalvez and Martinez, 2008). However, there is a substantial lack of quantitative evidence on investment behavior considering asset specificity, incomplete contracts and hold-up hazards.

Chapter I seeks to address this research gap in providing quantitative evidence on the consequences of different organizational structures and profit allocation mechanisms with respect to investment incentives.

An experimental analysis based on the theoretical framework of property rights theory (Grossman and Hart, 1986; Hart and Moore, 1990; Hart, 1995) is applied to the specific context of railways. In particular, three different organizational structures are modeled, i.e. vertical integration, vertical separation, and a hybrid form.

Economic theory predicts that vertical integration fosters socially optimal investment, whereas, due to potential hold-up problems, both vertical separation and hybrid forms cause severe underinvestment. These theoretical predictions are tested in a laboratory experiment and evidence suggests that, in a vertically integrated environment, the level of investment in rolling stock and in rail infrastructure is roughly socially optimal. However, the results do not confirm a clear discrepancy in results between vertical separation and the hybrid organizational structure, as predicted by model-theory. This contradiction might also be explained by the existence of social preferences.

2. The Costs and Benefits of Networking*

In the pursuit of optimizing the structure of the value creation process, externalizing certain activities has become a valid option for many firms, as is demonstrated by the increase of inter-firm alliances and related network forms of organization (Gulati, 1995). In particular, network forms of organization are emphasized to offer specific means of governance which allow for controlling exchange relations in a superior way, thereby leading to a competitive advantage of interconnected firms.

Yet, from a research perspective there is a striking imbalance of studies promoting the beneficial effects of networking as opposed to those emphasizing potential downsides of interfirm collaboration (e.g. Labianca and Brass, 2006; Parkhe, Wasserman and Ralston, 2006). Hence, the costs and dysfunctionalities of networking remain underexplored. Accordingly, Podolny and Page (1998: 66) state that the "attention to the functionality of network forms of organization explains why economic actors rely on network forms of organization, but it does not explain why they do not." Moreover, they argue that the widespread approach to concentrate on analyzing the benefits of networking might be useful for explaining varying reliance on networks *between* industries, since the relevance of collaboration-related benefits, such as learning and knowledge transfer, could differ from sector to sector. However, it still remains unclear why a considerable fraction of actors *within* an industry may constantly rely on hierarchical governance modes without investing in network building.

Chapter II seeks to address this issue in providing a brief conceptualization of the sources of networking costs and benefits based on findings from diverse theoretical perspectives such as transaction cost theory (Williamson, 1985), the resource based view (e.g. Barney, 1991) and social network theory (e.g. Granovetter, 2005; Uzzi, 1996). Important sources of benefits are knowledge transfer, relation-specific investments (e.g. in dedicated machinery or human capital), and access to complementary and/or scarce resources (e.g. reputation). Important costs

* This chapter originates from joint work with Thomas Ehrmann. An early version of this chapter was presented at the EMNet 2007 Conference in Rotterdam, NL. See: Ehrmann, T. and Schmale, H. (2007), The Costs and Benefits of Networking. Working Paper presented at the International Conference on Economics and Management of Networks, 2007, Rotterdam, NL.

involved in networking are transaction costs, originating from the need to control potentially opportunistic partners (e.g., White and Siu-Yun Lui, 2005; Williamson, 1996) and coordination costs related to coordination efforts independent from incentive conflicts (Camerer and Knez, 1997; Park and Ungson, 2001).

Moreover, and in line with findings from recent research on networking, non-linearities in the relationship between the level of networking activity and the costs and benefits of networking are considered. Hence, the argumentation explicitly accounts for varying value effects of different activity levels *within* the governance mode of networking since intensity of usage has been revealed as a crucial determinant of optimal institutional choice (Burt, 2005; Uzzi, 1997).

Finally, the consideration of opportunity costs – being defined as foregone net benefits from the most attractive alternative resource employment (e.g. Varian, 1997) – is a necessary precondition for quantifying the competitive advantage a particular governance form actually offers. The disregard of an opportunity cost calculus would impede the comparison of alternatively available governance modes, and consequently, the identification of competitive advantage. A simple heuristic model summarizes the main thoughts and illustrates that too extensive a reliance on network resources can imply a competitive disadvantage for interconnected firms.

3. The Impact of e-Word-of-Mouth on the Sales Distribution in Online Book Retailing*

The value creation potential of customers is receiving more and more recognition by practitioners and scholars. Firms more frequently "outsource" certain activities which were formerly assigned to the distributor role to customers. Examples are the implementation of automatic recommender systems and community-based product review platforms in online retailing. Via these systems, customers co-create value in reducing other potential customers' search costs. They do this either indirectly by just purchasing goods, i.e. filtering tools analyze product choices, infer taste patterns and recommend adequate products to like-minded customers (Smith, Bailey and Brynjolfsson, 1999), or directly by writing product reviews, thereby lowering product quality uncertainty. In this respect firms combine own resources and capabilities with those of the customer and thereby create new business opportunities.

The insight that customers are a critical source of value is the backbone of a recent phenomenon called "the long tail" (Anderson, 2006). In a nutshell, the long tail phenomenon describes the circumstance that in online retailing niche offerings account for a higher fraction of total sales than in offline retailing. This is driven by two facts. Firstly, in the online channel physical restrictions and cost constraints on assortment size disappear because of unlimited shelf space and centralized warehousing. Secondly, the above mentioned search cost reductions allow customers to venture into the niche to find products that better fulfill personal requirements than already known mainstream products. This in turn reflects one of the two basic assumptions which characterize the long tail literature. The assumption is that consumers differ in preferences, i.e. consumers value niche products which are tailored to their individual preferences more than products that are designed for mass appeal (e.g. Anderson, 2006; Elberse, 2008). The second assumption is that products differ in terms of search cost (Brynjolfsson, Hu and Simester, 2006). Specifically, key products which are heavily promoted by extensive marketing campaigns (e.g.

* This chapter originates from joint work with Thomas Ehrmann. An early version of this chapter was presented at the ICIS 2008 Conference in Paris, F. See: Ehrmann, T. and Schmale, H. (2008), The Hitchhiker's Guide to the Long Tail: The Influence of Online Reviews and Product Recommendations on Book Sales – Evidence from German Online Retailing. Proceedings of the International Conference on Information Systems, 2008, Paris, F.

TV ads, magazine promotion etc.) are visible per se. Niche products on the other hand are not visible per se and consequently profit more – in terms of sales – from a search cost reduction. From this follows that the shape of the demand curve will change as a result of a search cost reduction; from the rather steep sales distributions of bestseller-driven winner-take-all markets to more equally distributed sales in a world of "countless niches" (Anderson, 2006; Brynjolfsson, Hu and Smith, 2003; Brynjolfsson et al., 2006). The question now is whether this prediction holds and whether actors within the value chain should abandon bestseller strategies in favor of long tail strategies.

Despite the widespread belief among researchers and practitioners that user-generated online reviews and discussion forums are crucial instruments for driving demand down the tail there still is only sparse empirical evidence on the effectiveness of e-Word-of-Mouth (eWOM) instruments for reducing search costs in the long tail, and particularly, on their role in changing the sales distribution.

Exploring the long tail phenomenon, chapter III empirically analyzes whether online reviews, discussion forums, and product recommendations reduce search costs and if they can promote an altering of the sales distribution towards equality in online book retailing. By developing an innovative approach, the first long tail conversion model for the German online market is provided – based on publicly available sales data. A data set containing over 30,000 different books is analyzed and the results show that e-Word-of-Mouth reduces search costs by facilitating the identification of adequate books and their quality assessment. Moreover, unconditional quantile regressions reveal differences in the functionality of eWOM with respect to popular and unpopular books.

4. Buying Without Using: Biases of German BahnCard Buyers*

While data mining processes and thus value co-creation by customers is a core element of many digital businesses, firms in non-digital markets generally face substantial restrictions to the observability of customer behavior on an individual basis. For instance in the railway market, standard ticket purchases are not personalized and consequently transport providers can draw conclusions from purchase behavior on an aggregate level but not on an individual level. However, information on transport usage on an individual level can hold critical information for optimizing the value creation process in terms of answering basic questions like "what should be sold and how should it be sold?".

In an attempt to raise intermodal competitiveness of rail travel and furthermore to allow for segment-specific price differentiation for occasional and frequent travelers, the German rail company Deutsche Bahn AG introduced different two-part tariffs. These allow actors to pay a fixed up front fee and then receive a reduction on each ticket purchase (Dolan and Simon, 1996). The customer then has to find the contract that minimizes the price for every mile traveled.

A frequent approach to analyzing consumer behavior – when actual usage is unobserved – is to draw inferences from contract choice on preferences and thus demand (e.g. Miravete and Röller, 2003). A standard assumption then is that consumers have rational expectations about their future consumption frequency and choose the utility maximizing contract. However, if customers display non-standard preferences and beliefs, inferences made under the assumption of rational expectations can lead to substantial error (Della Vigna and Malmendier, 2006). A growing body of literature suggests that consumers indeed deviate from standard utility maximizing behavior in many cases and choose contracts that are sub-optimal from an ex post perspective considering actual usage (e.g. Della Vigna and Malmendier, 2006; Kling and van der Ploeg, 1990; Kridel, Lehman and Weisman, 1993; Lambrecht and Skiera, 2006; Miravete, 2002; Mitchel and Vogelsang, 1991; Train, Ben-Akiva and Atherton, 1989). Profit-maximizing firms should respond to

* This chapter originates from joint work with Thomas Ehrmann and Alexander Dilger.

these non-standard features of behavior in their product and contract design (Della Vigna and Malmendier, 2006).

Chapter IV of this thesis addresses this issue by using a large data set of German railway travelers for the first time to analyze the purchasing decision for fare-reducing BahnCards. The data comprise detailed information on customers' individual demographic characteristics, BahnCard contract choices and individual transaction data, i.e. ticket purchase behavior. The sample of more than 250,000 contracts was drawn from the population of members of the company's customer loyalty programs "bahn.bonus" and "bahn.comfort".

It is expected that tariff choice is neither completely rational nor irrational, but bounded-rational in a meaningful way. Actually the study predicts a flat-rate bias, i.e. many actors choose a contract involving a high fixed fee and low variable costs although another contract with a lower fixed fee and higher variable costs would have resulted in a lower billing rate, given their demonstrated demand (Nunes, 2000). Moreover, it is estimated that this bias is higher for more expensive BahnCards, while nevertheless the buyers of these BahnCards travel more than those of cheaper ones and renew their cards with a higher probability. Although many contracts are under-used, the average BahnCard should be worth its money, implying that the correctly using customers win more than the under-users loose or, alternatively, that the expected utility of any particular BahnCard is positive. Finally it is expected that learning exists and the usage of BahnCards improves over time, as well in the aggregate of all users as for individual ones. The empirical results approve these hypotheses for the most part, especially for the more expensive BahnCard50, whereas the under-use of the cheaper BahnCard25 is so extensive that it is not worthwhile on average.

IV. Integrative Framework

Figure 2 provides an overview of the systematization of this thesis's chapters. The first two chapters basically focus the trend of disintegration of production processes. Chapter I investigates the impact of a disintegration of the production process in railways on investment incentives thereby focusing the structure and governance dimensions of the value creation process. Chapter II analyzes the recent increase in interfirm collaborations from a theoretical perspective. In drawing attention to the costs of networking structural and governance aspects are focused. The next chapters, chapter III and chapter IV, turn to the downstream segments of the basic value chain in focusing on the value co-creation potential of customers and thus on the structure and content dimensions. Figure 3 provides a brief summary of research aims, methods and findings.

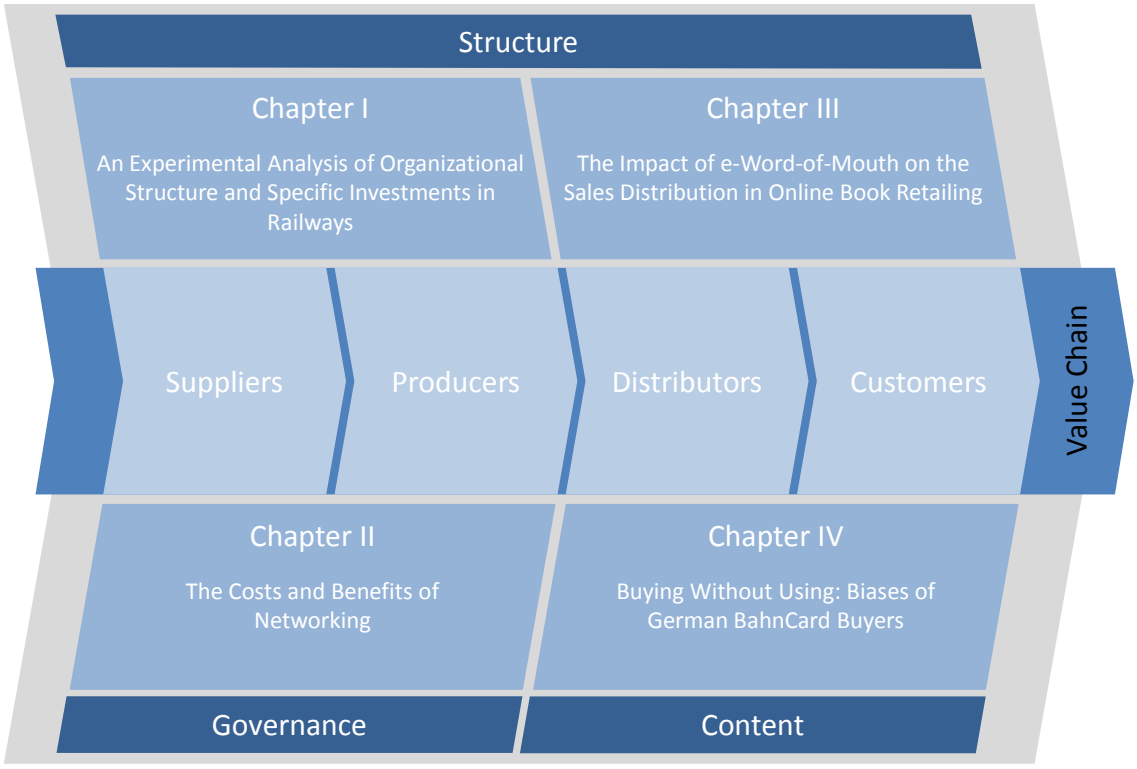

Figure 2. *Systematization of Chapters*

Chapter I:
An Experimental Analysis of Organizational Structure and Specific Investments in Railways

Research Aim:
Providing quantitative evidence on investment incentives under different organizational structures in railways, thereby considering incomplete contracts, asset specificities and hold-up hazards.

Methodological Approach:
Experimental analysis; three different organizational structures are investigated: vertical integration, separation and a hybrid form.

Results:
Vertical integration roughly yields socially optimal investment levels; vertical separation and the hybrid mode cause underinvestment on a nearly identical level, this is mainly driven by the relatively high investments in the separation case.

Chapter III:
The Impact of e-Word-of-Mouth on the Sales Distribution in Online Book Retailing

Research Aim:
Providing empirical evidence on the impact of customer reviews, discussion forums and product recommendations on search costs and on the distribution of sales in online book retailing.

Methodological Approach:
Empirical analysis using a sample of over 30,000 books; estimation of the first long tail conversion model for the German market using publicly available sales data; unconditional quantile regressions.

Results:
E-Word-of-Mouth reduces search costs by facilitating the identification of adequate books and their quality assessment. Unconditional quantile regressions reveal differences in the functionality of reviews, recommendations and discussions for popular and unpopular books.

Chapter II:
Costs and Benefits of Networking

Research Aim:
Drawing attention to the dysfunctionalities and costs of networking and providing a more balanced approach to the assessment of potential competitive advantages of interconnected firms considering non-linear developments of networking benefits and costs.

Methodological Approach:
Conceptualization of relevant findings from diverse theoretical perspectives such as TCE, the resource based view and social network theory.

Results:
Non-linearities in the relationship between the level of networking activity and the costs and benefits of networking can lead to a competitive disadvantage of interconnected firms at too high activity levels.

Chapter VI:
Buying Without Using: Biases of German BahnCard Customers

Research Aim:
Providing empirical evidence on tariff choice and usage behavior of German BahnCard customers; extending the literature on flat-rate biases.

Methodological Approach:
Empirical analysis of a large dataset containing over 250,000 contracts with detailed information on demographics, renewal behavior and individual transaction data, i.e. ticket purchases.

Results:
There exists a flat-rate bias, i.e. many customers underuse their contracts; the magnitude of the bias is higher for more expensive BahnCards while customers with higher BahnCards travel more and renew contracts with a higher probability; BahnCards are worthwhile on average with exception of the cheapest BahnCard; usage gets better over time.

Figure 3. *Research Intentions, Methodological Approaches and Results*

V. REFERENCES

Amit, R. and Zott, C. (2001). Value Creation in E-Business. *Strategic Management Journal*, 22: 493-520.

Anderson, C. (2006). *The Long Tail.* New York, NY: Hyperion.

Applegate, L. (2001). Emerging Networked Business Models: Lessons from the Field. Harvard Business School Note, 9-801-172.

Bakos, J.Y. and Brynjolfsson, E. (1997). Organizational Partnerships and the Virtual Corporation. In Kramer, C. (Ed.) *Information Technology and Industrial Competitiveness: How Information Technology Shapes Competition*. Boston, MA: Kluwer Academic Publishers.

Barney, J.B. (1991). Firm Resources and Sustained Competitive Advantage. *Journal of Management*, 17: 99-120.

Baum, J.A.C. and Oliver, C. (1991). Institutional Linkages and Organizational Mortality. *Administrative Science Quarterly,* 36: 187-218.

Bettis, R.A. and Hitt, M.A. (1995). The New Competitive Landscape. *Strategic Management Journal*, 16: 7-19.

Bitzan, J.D. (2003). Railroad Costs and Competition: The Implications of Introducing Competition to Railroad Networks. *Journal of Transport Economics and Policy*, 37: 201-225.

Brynjolfsson, E., Hu, Y.J. and Simester, D. (2006). Goodbye Pareto Principle, Hello Long Tail: The Effect of Search Costs on the Concentration of Product Sales. Working Paper, MIT Sloan School of Management, Cambridge, MA.

Brynjolfsson, Erik, Hu, Y.J. and Smith, M. (2003). Consumer Surplus in the Digital Economy: Estimating the Value of Increased Product Variety at Online Booksellers. *Management Science,* 49: 1580-1596.

Brynjolfsson, E., Malone, T., Gurbaxani, V. and Kambil, A. (1994). Does Information Technology Lead to Smaller Firms? *Management Science*, 40: 1628-1645.

Burt, R.S. (2005). *Brokerage and Closure: An Introduction into Social Capital.* New York: Oxford University Press.

Camerer, C. and Knez, M. (1997). Coordination in Organizations: A Game Theoretic Perspective. In Shapira, Z. (Ed.) *Organizational Decision Making.* Cambridge, MA: Cambridge University Press.

Chandler, A.D., Jr. (1977). *The Visible Hand: The Managerial Revolution in American Business.* Cambridge, MA: Belknap Press.

Ciborra, C.U. (1993). *Teams Markets and Systems*. Cambridge, MA: Cambridge University Press.

Cordella, A. (2006). Transaction Costs and Information Systems: Does IT Add Up? *Journal of Information Technology*, 21: 195-202.

Crocker, K.J. and Masten, S.E. (1991). Pretia Ex Machina? Prices and Process in Long-Term Contracts. *Journal of Law and Economics,* 34: 69-99.

DellaVigna, S. and Malmendier, U. (2006). Paying Not to Go to the Gym. *American Economic Review*, 96: 694-719.

Dolan, R.J. and Simon, H. (1996). *Power Pricing: How Managing Price Transforms the Bottom Line.* New York, NY: The Free Press.

Driessen, G., Lijesen, M. and Mulder, M. (2006). The Impact of Competition on Productive Efficiency in European Railways. CPB Discussion Paper, 71.

Ehrmann, T. (2003). Business Models in the New Economy. In Jones, D. (Ed.) *New Economy Handbook*. San Diego/London: Academic Press.

Elberse, A. (2008). Should You Invest in the Long Tail? *Harvard Business Review*, July/August: 88-96.

European Commission (1991). Council Directive 91/440/EEC of 29 July 1991 on the Development of the Community's Railways.

European Commission (1996). A Strategy for Revitalising the Community's Railways. White Paper. COM (96) 421 final, 30 July 1996.

European Commission (2001). Directive 2001/12/EC of the European Parliament and of the Council of 26 February 2001 Amending Council Directive 91/440/EEC on the Development of the Community's Railways.

Friebel, G., Ivaldi, M. and Vibes, C. (2003). Railway (De)Regulation: A European Efficiency Comparison. IDEI Report No. 3 on Passenger Rail Transport, University of Toulouse.

Granovetter, M. (2005). The Impact of Social Structure on Economic Outcomes. *Journal of Economic Perspectives,* 19: 33-50.

Grossman S. and Hart, O. (1986). The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration. *Journal of Political Economy*, 94: 691-719.

Growitsch, C. and Wetzel, H. (2009). Testing for Economics of Scope in European Railways: An Efficiency Analysis. *Journal of Transport Economics and Policy*, 43: 1-24.

Gulati, R. (1995). Does Familiarity Breed Trust? The Implications of Repeated Ties for Contractual Choice in Alliances. *Academy of Management Journal*, 38: 85-112.

Hamel, G. (1991). Competition for Competence and Interpartner Learning within International Strategic Alliances. *Strategic Management Journal,* Summer Special Issue, 12: 83-103.

Hart, O. (1995). *Firms, Contracts and Financial Structure*. New York, NY: Oxford University Press.

Hart O. and Moore, J. (1988). Incomplete Contracts and Renegotiation. *Econometrica,* 56: 755-785.

Hitt, M.A., Keats, B. and DeMarie, S.M. (1998). Navigating in the New Competitive Landscape: Building Strategic Flexibility and Competitive Advantage in the 21st Century. *Academy of Management Executive*, 12: 22-41.

Hoskisson, R.E., Hitt, M.A., Wan, W.P. and Yiu, D. (1999). Theory and Research in Strategic Management: Swings of a Pendulum. *Journal of Management*, 25: 417-456.

Joskow, P.L. (2003). Vertical Integration. In Menard, C. and Shirley, M.M. (Eds.) *Handbook of New Institutional Economics*. Dordrecht: Springer.

Katz, M.L. and Shapiro, C. (1985). Network Externalities, Competition and Compatability. *American Economic Review*, 75: 424-440.

Klein, B. Crawford, R. and Alchian, A. (1978). Vertical Integration, Appropriable Rents, and the Competitive Contracting Process. *Journal of Law and Economics,* 21: 297-326.

Kling, J.P. and van der Ploeg, S.S. (1990). Estimating Local Elasticities with a Model of Stochastic Class of Service and Usage Choice. In Fontenay, A. de, Shugard, M.H. and Sibley, D.S. (Eds.) *Telecommunications Demand Modeling. An Integrated View.* Amsterdam: North Holland.

Kridel, D.J., Lehman, D.E. and Weisman, D.L. (1993). Option Value, Telecommunication Demand, and Policy. *Information Economics and Policy*, 5: 125-144.

Labianca, G. and Brass, D.J. (2006). Exploring the Social Ledger: Negative Relationships and Negative Asymmetry in Social Networks in Organizations. *Academy of Management Review,* 31: 596-614.

Lambrecht, A. and Skiera, B. (2006). Paying Too Much and Being Happy About It: Existence, Causes and Consequences of Tariff-Choice Biases. *Journal of Marketing Research*, 18: 212-223.

Langlois, R.N. (2003). The Vanishing Hand: The Changing Dynamics of Industrial Capitalism. *Industrial and Corporate Change*, 12: 351-385.

Malone, T.W., Yates, J. and Benjamin, R.I. (1987). Electronic Markets and Electronic Hierarchies: Effects of Information Technology on Market Structure and Corporate Strategies. *Communications of the ACM*, 30: 484-497.

Miravete, E.J. (2002). Estimating Demand for Local Telephone Service with Asymmetric Information and Optional Calling Plans. *Review of Economic Studies*, 69: 943-971.

Miravete, E.J. and Röller, L.-H. (2003). Competitive Non-Linear Pricing in Duopoly Equilibrium: The Early US Cellular Telephone Industry. CEPR Discussion Papers: No. 4069.

Mitchell, B.M. and Vogelsang, I. (1991). *Telecommunication Pricing: Theory and Practice.* Cambridge, UK: Cambridge University Press.

Nash, C.A. and Preston, J.M. (1994). Competition in Rail Transport: A New Opportunity for Railways. Working Paper 397, Institute for Transport Studies, University of Leeds.

Nash, C.A. and Rivera-Trujillo, C. (2007). Rail Reform in Europe: Issues and Research Needs. In Rietveld, P. and Stough, R.R. (Eds.) *Institutions and Sustainable Transport: Regulatory*

Reforms in Advanced Economies. Cheltenham, UK and Northampton, MA, USA: Edward Elgar.

Norman, R. and Ramirez, R. (1993). From Value Chain to Value Constellation: Designing Interactive Strategy. *Harvard Business Review*, 71: 65-77.

Nunes, J.C. (2000). A Cognitive Model of People's Usage Estimations. *Journal of Marketing Research,* 37: 397-409.

Park, S.H. and Ungson, G.G. (2001). Interfirm Rivalry and Managerial Complexity: A Conceptual Framework of Alliance Failure. *Organization Science,* 12: 37-53.

Parkhe, A., Wasserman, S. and Ralston, D.A. (2006). New Frontiers in Network Theory Development. *Academy of Management Review,* 31: 560-568.

Pillai, K.G. (2006). Networks and Competitive Advantage: A Synthesis and Extension. *Journal of Strategic Marketing,* 14: 129-145.

Pittman, R.W. (2007). Make or Buy on the Russian Railway? Coase, Williamson, and Tsar Nicholas II. *Economic Change and Restructuring*, 40: 207-21.

Podolny, K.L. and Page, J.M. (1998). Network Forms of Organization. *American Review of Sociology,* 24: 57-76.

Powell, W.W., Koput, K.W. and Smith-Doerr, L. (1996). Interorganizational Collaboration and the Locus of Innovation: Networks of Learning in Biotechnology. *Administrative Science Quarterly,* 41: 116-145.

Prahalad, C.K. and Ramaswamy, V. (2004). Co-Creation Experiences: The Next Practice in Value Creation. *Journal of Interactive Marketing*, 18: 5-14.

Preston, K. (1996). The Economics of British Rail Privatisation: An Assessment. *Transport Reviews*, 16: 1-21.

Rothengatter, W. (2001). How Good is First Best? Marginal Cost and other Pricing Principles for User Charging in Transport. Working Paper, University of Karlsruhe.

Sanchez, P.C., Monsalvez, J.M.P. and Martinez, L.S. (2008). Vertical and Horizontal Separation in the European Railway Sector: Effects on Productivity. Documentos de Trabajo, Fundacion BBVA, 12.

Smith, A. (1776). *An Inquiry into the Nature and Causes of the Wealth of Nations*. Edited by Sálvio M. Soares. MetaLibri, 2007, v.1.0p.

Smith, M.D., Bailey, J. and Brynjolfsson, E. (1999). Understanding Digital Markets: Review and Assessment. In Brynjolfsson, E. and Kahin, B. (Eds.) *Understanding the Digital Economy*. Cambridge, MA: MIT Press.

Stigler, G.J. (1951). The Division of Labor is Limited by the Extent of the Market. *Journal of Political Economy*, 59: 185-193.

Teece, D.J. (1987). Profiting From Technological Innovations: Implications for Integration, Collaboration, Licensing, and Public Policy. In Teece, D.J. (Ed.) *The Competitive Challenge: Strategies for Industrial Innovation and Renewal*. Cambridge, MA: Ballinger.

Train, K.E., Ben-Akiva, M. and Atherton, T. (1989). Consumption Patterns and Self-Selecting Tariffs. *Review of Economic and Statistics,* 1: 62-73.

Uzzi, B. (1996). The Sources and Consequences of Embeddedness for the Economic Performance of Organizations: The Network Effect. *American Sociological Review,* 61: 674-698.

Uzzi, B. (1997). Social Structure and Competition in Interfirm Networks: The Paradox of Embeddedness. *Administrative Science Quarterly,* 42: 35-67.

Varian, H.R. (1997). *Intermediate Microeconomics – Fourth Edition.* New York: Norton.

White, S. and Siu-Yun Lui, S. (2005). Distinguishing Costs of Cooperation and Control in Alliances. *Strategic Management Journal,* 26: 913-932.

Williamson, O.E. (1975). *Markets and Hierarchies: Analysis and Antitrust Implications.* New York: Free Press.

Williamson, O.E. (1985). *The Economic Institutions of Capitalism.* New York: Free Press.

Williamson, O.E. (1996). *The Mechanisms of Governance.* New York: Oxford University Press.

PART B

I. AN EXPERIMENTAL ANALYSIS OF ORGANIZATIONAL STRUCTURE AND SPECIFIC INVESTMENTS IN RAILWAYS*

1. Abstract

We analyze the impact of different organizational structures on incentives to invest in railways: vertical integration, vertical separation, and a hybrid form. Economic theory predicts that vertical integration fosters socially optimal investment, whereas, due to potential hold-up problems, both vertical separation and hybrid forms cause severe underinvestment. We test these theoretical predictions in a laboratory experiment and find evidence that, in a vertically integrated environment, the level of investment in rolling stock and in rail infrastructure is roughly socially optimal. The absence of a clear discrepancy in our experimental results between vertical separation and the hybrid organizational structure, contradicting the predictions of model-theory, is surprising and can be attributed to the relatively high investments in the separated model. This contradiction might also be explained by the existence of social preferences.

* This chapter originates from joint work with Thomas Ehrmann, Karl-Hans Hartwig and Torsten Marner. For the discussion paper see: Ehrmann, T., Hartwig K.-H., Marner, T. and Schmale, H. (2009), Specific Investments and Ownership Structures in Railways: An Experimental Analysis. Discussion Paper No. 12 of the Institute of Transport Economics, University of Muenster.

2. Introduction

"However beautiful the strategy, you should occasionally look at the results."
Winston Churchill

Since European policy is presently demanding more competition in European railways, vertical relationships in railways are subject of substantial controversy. Advocates of a vertical separation of infrastructure from transport operation argue that, even if vertically integrated firms are obliged to grant third-party access to railway infrastructure, potential for market foreclosure and discrimination will continue to exist and competition will remain restricted (Nash and Preston, 1994; European Commission, 1996; Link, 2003). Therefore, vertical separation is regarded as the only way to enhance competition within the railway industry. Proponents of vertical integration argue that an institutional separation would reduce economic welfare, because of losses of economies of scope, of lower consumer attractiveness due to coordination failure and of insufficient investment as a result of asset specificity, incomplete contracts and hold-up hazards (Cantos, 2001; Pfund, 2003; Pittman, 2007).

There is a large body of *theoretical* research that supports these positions, highlighting the drawbacks and inefficiencies of vertical separation, as well as the advantages and disadvantages arising from vertical integration solutions. This abundance of theoretical research contrasts with the low number of recent empirical studies, even though a very substantial number of institutional settings for the railway sector have been established worldwide in the mean-time (Gomez-Ibanez, 2004; Cantos and Campos, 2005; Nash, 2006). According to Nash and Rivera-Trujillo (2007) the lack of *empirical* evidence is due to the fact that most of these settings are not comparable, in particular due to the short time horizon of reform experiences. Sufficiently large data bases necessary for analytical research are available only for a few of the relevant variables. Most previous studies on vertical separation test for economies of scope (Preston, 1996; Growitsch and Wetzel, 2009) and analyze the implications on competition and on efficiency and productivity growth (Nash and Preston, 1994; Bitzan, 2003; Friebel et al., 2003; Driessen et al., 2006; Sanchez et al., 2008). Up to now, there is only one empirical study on asset specificity, incomplete contracts, hold-ups and investment behavior in different institu-

tional settings of rail industry by Merkert, Nash and Smith (2008). From the perspective of the New Institutional Economics, the authors analyze the impact of the governance structures of British, German and Swedish railways on competition and on the transaction costs of different interactions between infrastructure managers and train operators. The data have been collected by reviewing policy documents and contracts from seven pre-specified transaction areas and interviews with infrastructure managers, senior managers from train operators, regulators and industry associations. The results show that, although asset specificity and incomplete contracts do exist, the frequency, uncertainty and complexity of coordination and contractual interactions are perceived as more relevant than investment hold-up or lock-in issues (p. 27). All in all, the authors conclude that vertical separation turns out to be the "clearest approach in terms of non-discrimination" and viable at reasonable cost, in "terms of transaction cost economics" (p. 40).

In order to determine whether a separate railway organization would reduce or even eliminate the incentive to invest on one or even both sides (the infrastructure provider and the transport operator), so that underinvestment may occur, raising costs and diminishing welfare in the long run, we adopt an experimental approach. Based on the seminal work of Grossman and Hart (1986), Hart and Moore (1990) and Hart (1995) on specific investments and the structure of vertical relationships, we model investment behavior in various institutional arrangements in railways, hypothesize corresponding investment levels and test these hypotheses through experimental research. Such research provides an alternative framework to systematically designing varying institutional settings and analyzing the resultant incentive structures and their impact on economic behavior (Roth, 1995). In our case, the approach sheds some empirical light, from another perspective, on an important aspect of restructuring the European railway industry, an issue has so far been discussed by means of more or less qualitative arguments in case studies. The fundamental question is whether the investment incentives associated with a separate institutional arrangement can cause a long-term investment problem and welfare losses.

The structure of this chapter is as follows. In Section 3, we generally describe the hold-up problem and show the relevance of specific investments, opportunistic behavior and incomplete

contracts in railways. In Section 4, our model of investment behavior in different organizational structures is introduced, so that the expected investment levels can be hypothesized. Section 5 contains the experimental design. In Section 6, the experimental results with respect to investment incentives are presented. In Section 7, the results are discussed, before we close with a summary and some future perspectives.

3. Specific Investments, Incomplete Contracts and Underinvestment in Railways: Theory and Previous Research

According to modern institutional economics, a vertical separation of infrastructure from operations in network industries is an inferior form of organization, if investments are specific and contracts incomplete (Williamson, 1975 and 1985; Klein et al., 1978; Crocker and Masten, 1991). Because specific investments yield significantly lower values or lower gains from trade, when employed in a transaction other than originally intended (Joskow, 2003), the investor bears the risk of being exploited by an opportunistic transaction partner, who will appropriate the difference between the value of the investment in its first and second-best use (quasi-rent) in an ex-post bargaining process. If the investor anticipates the risk of a hold-up and if contractual arrangements to avert hold-ups are hindered by incomplete contracts, he will not undertake the investment at all. In network industries, underinvestment may occur on either side of the transaction. The vertical integration of infrastructure and operations could constitute an institutional setting which prevents disincentives to invest.

Previous research on the rail industry had indeed identified asset specificity in the network infrastructure as well as in the rolling stock. Primary arguments are the strong technical interdependency of both of the input factors and the fact that investments not only require significant financial resources, but most often are completely irreversible (Rothengatter, 2001; Gomez-Ibanez, 2004; Cantos and Campos, 2005; Pittman, 2005). Various empirical studies have attempted to document and estimate asset specificity. Yvrande-Billon (2004) estimated a high level of specificity, measured by the impossibility of re-deployment of the rolling stock of British railways. According to Affuso and Newbery (2002), up to 82 per cent of each asset of the transport companies in Great Britain are specific. Ferreira (1997), Crozet (2004), Bouf et al. (2005), von Hirschhausen and Siegmann (2004) and Merkert et al. (2008) detected asset specificities of different kinds (physical specificity, site specificity, dedicated specificity and temporal specificity) and different levels of relevance upstream on the infrastructure level as well as downstream on the operational level of rolling stock.

Common examples of asset specificity in the railway context are investments in high-speed rail lines and in modern signal and safety technology, such as the European Train Control System (ETCS) which allows for higher capacities and higher operating densities, due to the economization of permanent signaling equipment through the direct transmission of propulsion command via GSM (de Rus and Nombela, 2007; International Union of Railways, 2003). Investment in high-speed rail tracks and ETCS-infrastructure is enormously cost-intensive for infrastructure companies and requires simultaneous investments in rolling stock by a transport operator who pays for and uses the track for a period sufficient for an amortization of the infrastructure investment. As Pittman (2007) points out, "a track operator can make certain investments to improve efficiency and performance, but the realization of these benefits depends significantly on actions taken by the train operator". If the technological demand for transport operation exhibits a lower level than high-speed rail and ETCS, it is impossible to maintain an appropriate price for track usage (Nash, 2005). The value of the investments will decrease. This is correspondingly true vice versa, if the transport operator invests in high-speed trains or implements ETCS in rolling stock without corresponding investments from the infrastructure operator, as in the case of Virgin Rail, a British transport company (Pfund, 2003).

In order to avoid the hazards of "downgrading" the infrastructure and rolling stock, investments have to be coordinated very exactly, so as to produce the final output of transportation and to improve quality in terms of timesaving and safer transport. With separate environments for rail infrastructure and transport operation, efficient coordination fails to take place, because of disincentives to invest on both sides as a result of hold-up risks and incomplete contracts. Since investment behavior is neither fully observable nor enforceable by law "subject to shirking and opportunism, the investments on both sides may not be made and economic welfare will suffer as a result" (Pittman, 2007). A potentially superior institutional solution could take the form of vertical integration, which is proved empirically by the experiment described in the present chapter.

4. Theoretical Model

4.1 Basic Structure

In order to test the hypothesis that vertical integration is the superior form of organization in the railway industry with respect to asset specificities, we use the standard models of Grossman and Hart (1986), Hart and Moore (1990) and Hart (1995). Theory states that incentives to invest depend on ownership structure. Because ownership structures in European railways are established politically in conformity with EU guidelines (European Commission, 1991; 2001), we develop three different scenarios with exogenously predetermined ownership structures:

1. vertical separation
2. vertical integration
3. hybrid form of organization.

Ownership rights affect the incentive structure in determining the extent to which an investor can claim a generated surplus and thereby recover at least his *initial* investment costs. Hence, ownership creates the incentive to invest. We assume that investments are embodied in physical capital such as rolling stock and rail infrastructure, rather than in human capital. From this, it follows that the value of the investment is not bound to the investor, but solely to the respective asset.[8] Furthermore, ownership of an asset assigns the right to make an investment, as well as the ability to transfer this right, since making the investment is assumed not to be specific to a particular individual.

Investment decision rights are allocated to the transport operator (F_1) and the infrastructure operator (F_2), together producing the final good of railway transport by a combination of the two specific assets of rolling stock (a_1) and rail infrastructure (a_2). In fact, the transport operator uses

[8] If the investments were embodied in human capital, rather than in physical capital, an acquisition of the complementary asset, that is vertical integration, would not enable the new owner to generate a full surplus, because part of the investment's value would be tied to the former owner himself. Thus, in the case of integration, the acquiring firm would still have to negotiate with the former owner, in order to obtain full access to the investment, although it already controls the physical asset.

the track to produce transport activities. Depending on the ownership structure, both actors can either be completely autonomous firms or departments within one integrated firm. The gross surplus derived from the transport activity (S) depends on specific investments in rolling stock (i_1) and in the rail infrastructure (i_2): $S(i_1,i_2)$. Investments increase the productivity of the assets and are made in period $t = 1$, in which investment costs $c(i_1)$ and $c(i_2)$ accrue to the investing party. Although in $t = 1$, it is clear that specific investments are required to produce the final good of transportation, uncertainty prevails as to the precise asset configuration. This is due to the fact that, particularly in the context of railways, the production of the final good is highly complex and therefore, the costs of defining a comprehensive contract over the exact uses of a_1 and a_2 are assumed to be prohibitively high. This uncertainty also means that ex-ante contracting involving the division of the surplus from cooperation is not feasible. Hence, the allocation of gross surplus cannot take place until investment is sunk and uncertainty is resolved in the next period, $t = 2$. Figure 4 summarizes this chronology of action.

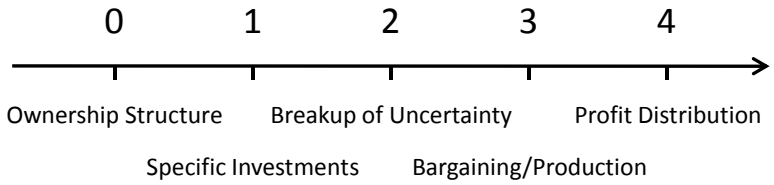

Figure 4. *Chronology of Action*

We assume that in the case considered here, the gross surplus from production $S(i_1,i_2)$ is defined as $S(i_1,i_2) = 20(i_1 + i_2)$. Investment costs $c(i_1)$ and $c(i_2)$ are defined as $c(i_1) = 15i_1$ and $c(i_2) = 15i_2$. Investments can be chosen from the interval $i_1, i_2 \in \{1,2,...,10\}$.

4.2 Vertical Separation

In the case of vertical separation, one of the pure forms of privatization alternatives, the transport operator F_1 and the infrastructure operator F_2 are completely autonomous firms. Each possesses one productive asset, that is, F_1 owns and controls a_1 and F_2 owns and controls a_2, so that the transport operator contributes to the provision of transport by making the rolling stock available, while the infrastructure operator contributes to the production of the final good of transportation, by providing the rail infrastructure. Consequently, both actors independently and simultaneously choose investments in $t = 1$. After uncertainty is resolved in $t = 2$, they bargain over the infrastructure charge and type, determining the division of the resultant gross surplus $S(i_1,i_2)$. Finally, when the actors reach agreement and trade occurs, ex-post pay pay-offs are realized for the transport operator (Π_1) and the infrastructure operator (Π_2), given by the following equations:

$$(1) \qquad \Pi_1 = \begin{cases} b \cdot S(i_1,i_2) - c(i_1) & \text{contractual agreement,} \\ -c(i_1) & \text{non-agreement.} \end{cases}$$

$$(2) \qquad \Pi_2 = \begin{cases} (1-b) \cdot S(i_1,i_2) - c(i_2) & \text{contractual agreement,} \\ -c(i_2) & \text{non-agreement.} \end{cases}$$

b denotes the transport operator's negotiated share of the surplus, $(1 - b)$ the infrastructure operator's share in the case of agreement. However, it is important to note that a surplus is generated only if both actors agree to trade. Otherwise, the production of the final good is impeded, since both actors withdraw their asset from the production process. Since investments are sunk, each has to bear his individual investment costs.

In our model, bargaining follows a Rubinstein alternating-offer structure with a maximum of ten bargaining rounds (Rubinstein, 1982) and a multiple-pie finite-horizon bargaining setting

(Sloof, 2004).[9] In each bargaining round, one round-pie is negotiated between the two players. The size of each round-pie is 1/10 $S(i_1,i_2)$. Both actors alternate in making offers with respect to the division of the ten round-pies, with the first offer being randomly assigned to one of the players. Actors are allowed to respond to offers in three different ways. Firstly, they can accept the offer and the round-pie of the current bargaining round, with all remaining round-pies being divided according to the agreement. Secondly, the responder can reject the offer and terminate negotiations. In this case, the current and all remaining round-pies are irrevocably lost and, consequently, both players receive nothing. Thirdly, the player can reject and submit a counter offer instead. Bargaining then proceeds to the next round and the current round-pie is lost, this in turn reflecting the cost of negotiation. Finally, b and $(1 - b)$ are determined by bargaining.

We assume that the gains from trade are divided according to the Nash bargaining solution, that is, a 50/50 division of the surplus (Nash, 1950), so that investments result from the optimization of equations (3) and (4):

$$\Pi_1 = \frac{1}{2} \cdot S(i_1, i_2) - c(i_1), \tag{3}$$

$$\Pi_2 = \frac{1}{2} \cdot S(i_1, i_2) - c(i_2).^{10} \tag{4}$$

Since $S(i_1,i_2) > c(i_1) + c(i_2)$ does apply, in a first-best world, where coordination between the two parties is feasible, F_1 as well as F_2 would have an incentive to invest the maximum amount of $i_{1,2}=i^{max}$. In the absence of hold-up threats, the parties could redistribute any increase in value by means of ex ante lump-sum transfers.

However, $1/2 \cdot (\partial S / \partial i_1) < \partial c_1$ and $1/2 \cdot (\partial S / \partial i_2) < \partial c_2$ imply that individually, in an incomplete contracting world with rational and self-interested actors, investment entails strictly nega-

[9] In the interest of simplification, we disregard any other discounting effects. This guarantees the implementation of an exactly symmetrical Nash bargaining solution (Nash, 1950). Any further discounting would have caused a first-mover advantage for the subject with the right of first offer. Here, backward induction predicts a sub-game perfect equilibrium with an equal share in the first bargaining round.

[10] This is true for $\partial S / \partial i > 0$ and $\partial^2 S / \partial^2 i = 0$. For the sake of simplicity, investment costs are assumed to be linear, so that $\partial c / \partial i > 0$ and $\partial^2 c / \partial^2 i = 0$.

tive net pay-offs. This follows from the fact that any increase in value, $\partial S / \partial i$, must be shared equally with the other partner, whereas increasing investment costs are incurred on one's own. Consequently, the marginal costs of investment exceed the marginal benefits. Both the transport operator and the infrastructure operator will invest the minimum of $i_1 = i_2 = i^{min}$, both anticipating opportunistic behavior in the form of a hold-up by the other party.[11] Given the abovementioned parameterization, $\Pi_1 = 10i_2 - 5i_1$ and $\Pi_2 = 10i_1 - 5i_2$ describe the individual optimization problems. Hence, for the transport operator, the choice of $i_1 = i_{min} = 1$ is optimal and for the infrastructure operator, it is optimal to choose $i_2 = i_{min} = 1$. This situation resembles a prisoners' dilemma and results in bilateral underinvestment. The prisoners' dilemma is documented by Figure 5 and depicts the profits accruing to the players at three different levels of investment (1, 5, 10). Combination (i_1, i_2) = (1; 1) is a Nash equilibrium with a resulting overall profit of $\prod_1+\prod_2=5+5=10$, which is inferior to individually unstable investments of 10, generating an overall profit of $\prod_1+\prod_2=50+50=100$.

i_1 ; i_2	**1**	**5**	**10**
1	5 ; 5	45 ; -15	95 ; -40
5	-15 ; 45	25 ; 25	75 ; 0
10	-40 ; 95	0 ; 75	50 ; 50

Figure 5. *Vertical Separation and Prisoners' Dilemma*

[11] Only after their own investment is sunk, do the agents learn of the other party's investment. In this respect, other constellations are also possible. Nöldeke and Schmidt (1998) and Smirnov and Wait (2004) concentrate, for example, on the problem of underinvestment in the case of sequential investments. In the railway context, existing monitoring and contract-enforcement problems imply that the application of simultaneous investments is advisable.

4.3 Vertical Integration

In the model of vertical integration, one fully integrated railway company owns the rights of control over the net infrastructure and the rolling stock, as well as the investment rights. Therefore, there is no hold-up hazard and the investor can fully internalize revenue derived from his investments. The integrated firm's investment incentives, therefore, are expressed in optimizing the following equation:[12]

$$(5) \qquad \Pi_2 = S(i_1, i_2) - c(i_1) - c(i_2).$$

Accordingly, the model predicts maximal investment levels for the integrated case: $i_1 = i_2 = i^{max}$. Given the abovementioned parameterization, $\Pi_2 = 5(i_1 + i_2)$ is true, investments in rolling stock and in transport operation reach their maximum levels: $i_1 = i_2 = i_{max} = 10$, and thus, $i_1 + i_2 = 20$.

In the separate structure, the need to recover investment costs implies an incentive to fully utilize the enhanced productivity of assets. This holds, because any quality loss due to a reduction in effort would be counterproductive, given one's own prior investment decision.[13] However, in an integrated arrangement, this is not true for all production steps, taking into account that the transport division cannot make investment decisions and does not have to bear any investment costs. Given individual self-interest and rational behavior, the result is shirking by the transport division, in order to minimize the disutility of work (Alchian and Demsetz, 1972). Although shirking reduces the efficiency of those parts of the production process involving the transport operator's effort and decreases the marginal benefits of investment, shirking does not affect investment incentives, because marginal benefits still significantly exceed marginal cost.

[12] Equation (5) denotes the integrated firms' ex-post pay-off in the case of Type 2 integration (see Hart, 1995: 35), that is, F_2 integrates F_1 and thus, becomes the sole owner of the entire set of assets, a_1 and a_2. Alternatively, we could have considered the case of Type 1 integration. Given the symmetry of parameterization, this would not have any effect on the theoretical predictions regarding investment incentives. Therefore, we content ourselves with the analysis of one case.

[13] Taking this aspect into consideration further strengthens the consistency of our experimental investigation, since this aspect implies interaction between F_1 and F_2, although F_2 is the sole owner of the assets and makes the investment decisions.

4.4 Hybrid Model

Considering the ownership-structure continuum, which is bounded by full integration on one side and full separation on the other, there are obviously several alternative hybrid organizational designs (Ferreira, 1997). Below, we model a structure which partly separates the rights to control the assets from those of making an investment. Accordingly, the final good of rail transportation is produced by a company which is subdivided into two divisions: the transport operator and the dominant infrastructure operator. The rights of control over rolling stock a_1 and rail infrastructure a_2 are assigned to the dominant infrastructure operator F_2.[14] Thus, analogously to the integrated case, this model also allows a full internalization of investments revenue. Nevertheless, this model differs in terms of investment-rights allocation. In particular, the dominating infrastructure operator transfers the responsibility and right to invest in the rolling stock to the dominated transport operator and pays remuneration w to F_1, after receiving the complete gross surplus from the investments. Hence, in the first step, the transport operator decides on investment i_1 and in the second step, he receives the compensation.

$$\Pi_1 = w - c(i_1), \tag{6}$$

$$\Pi_2 = S(i_1, i_2) - c(i_2) - w. \tag{7}$$

Because the infrastructure operator possesses all control rights and, therefore, full residual claimant status, his incentive is to invest the maximum amount $i_2 = i^{max}$. His marginal benefits from investment exceed the marginal costs. The transport operator only decides to invest $i_1 = i^{min}$ in anticipation of rational and self-interested behavior from the infrastructure operator, which implies the minimum compensation $w = w^{min}$=15.[15]

Given the above parameterization, $\Pi_1 = w - 15i_1$ and $\Pi_2 = 20i_1 + 5i_2 - w$ are true. In equilibrium, the transport operator invests an amount of $i_1 = i_{min} = 1$ and the infrastructure operator an amount of $i_2 = i_{max} = 10$. As a result, unilateral underinvestment occurs and the model of hybrid

[14] To ensure consistency with regard to the integration model, we obviously consider Type 2 integration here as well.
[15] Note that F_1 must invest at least the minimum of $i_1 = 1$.

organization is therefore inferior to the integrated model but, in terms of investment incentives, superior when compared to the separated model. Table 1 summarizes the standard theoretical predictions of investment behavior and bargaining outcomes, given our parameterization.

	Vertical Separation	Vertical Integration	Hybrid Model
Stage 1			
i_1	1	10	1
i_2	1	10	10
$i_{12} = i_1 + i_2$	2	20	11
Stage 2			
b	0,5		
w			15
No. of rounds	1	1*	1*

Asterisks indicate that the number of bargaining rounds is fixed due to the setup of treatments.

Table 1. *Summary of Hypotheses*

5. Experimental Design

The experimental investigation consisted of six treatments, both multi- and single-period games for each of the three models. The computer-based experiments were carried out at the Department of Economic Studies, Muenster University, Germany, in 2006. 256 respondents were recruited from a homogenous group of students studying economics at an advanced level. For each treatment, the subject group was divided into two subgroups, half being assigned the role of transport operator (F_1) and the other the role of infrastructure operator (F_2). Subjects kept their role throughout the experiment. They were paid according to performance and earned, on average, €12.50 per hour.

The multi-period treatment consisted of 12 rounds, during which each of the respondents were randomly and anonymously matched pair-wise. Instructions were handed out and read to all subjects. Thus, all had identical information about the rules and structure of the game. In order to ensure that any player knew the consequences of his decisions, we provided a simulation device. The simulator enabled the calculation of outcomes of investment and bargaining decisions throughout the game. The subjects were also informed about the matching procedure. Hence, reputational effects should be minimized. Furthermore, the subgroups were located in separate rooms and communication within the subgroup was strictly forbidden. This ensured that no player could forecast his current partner's decisions on the basis of past behavior.

Each round involved two stages, an investment and a bargaining stage. This applies to the single-period scenarios as well as to the multi-period scenarios.[16] The single-period games involved a single play of the two-stage game. At the investment stage, the subjects were requested to simultaneously choose their investment levels. However, as described above, the right to invest depends on the ownership structure. In the separated case and the hybrid case, both players – the one assigned the role of infrastructure operator and the counterpart – were provided with the right to invest in one of the assets. However, in the integrated case, the infrastructure

[16] For simplicity, we will call the second stage the *bargaining stage* throughout the remainder of the study, even though, in the integrated case and the hybrid case, the second stage is more likely to resemble a dictator game rather than a bargaining game.

operator had the right to invest in both of the assets, i.e. rolling stock and infrastructure. At the bargaining stage, the impact of the considered ownership structures on the nature of the game was greater than at the first stage. In the integrated and hybrid cases, stage two consisted of the choice of compensation level by the infrastructure player. At stage two of the separation scenario, the players were asked to allocate the surplus generated in stage one by the abovementioned bargaining process.

In each of the considered treatments, the subjects were informed about the chosen investment levels before proceeding to the bargaining stage. At the end of stage two of each treatment, payments were made according to the bargaining results and, finally, investment costs were incurred by the investors. In the multi-period games, both players subsequently moved to the next game period and were again randomly matched with a partner. In the single-period games, the experiment finished at this point and the subjects were paid according to their performance.

6. Experimental Results

In the multi-period treatments 44 subjects participated in the separated case, 44 in the hybrid case and 42 in the integrated case. With respect to investment behavior, the first result from the experiment is as follows:

Result 1. Average investment levels are maximized in the integration model. Integration induces levels of investments close to the social optimum.

The average total investment in rolling stock and rail infrastructure amounted to 14.70 in the separated case, 13.65 in the hybrid case, and 19.56 in the integration case. The experimental results thus evidently support the theoretical prediction with respect to the vertical integration model. In more than 92 per cent of the cases (466 out of 504), the respondents chose the efficient investment level $i_{max} = 10$.

Average Investment Levels				Mann-Whitney-U-Test					
	Integration	Hybrid	Separation	Int. and Sep.		Int. and Hyb.		Hyb. and Sep.	
Variable	Case	Case	Case	Z	p > Z	Z	p > Z	Z	p > Z
i_1	9.78	4.39*	7.51	-11.08	.000	-17.45	.000	-10.20	.000
i_2	9.79	9.26*	7.19	-11.39	.000	-4.64	.000	-8.19	.000
i_1+i_2	19.56	13.65	14.70	-14.87	.000	-17.41	.000	-3.14	.002

Asterisks indicate significant differences (<1%) within the same column, that is differences in investments between F_1 and F_2 in the respective treatment.

Table 2. *Investment Incentives, Two Sample, Non-Parametric Pair-Wise Tests*

A pair-wise comparison of the investments, conducted by means of the Mann-Whitney-U-Test, indicates the superiority of the integrated model ($p < .01$). Table 2 shows the results from three non-parametric tests, examining total investments for rounds 1-12.

Result 2. The hybrid case does not provide higher investment incentives than the separated model. Both the separation and the hybrid cases cause underinvestment with respect to the social optimum, but exceed theoretical predictions from the models.

A comparison of theoretical predictions with the investment results from the experiment reveals that, in the separation case as well as in the hybrid case, investments clearly exceed the equilibrium results of the models (separation case: 14,70 > 2, hybrid case: 13,65 > 11). The absence of clear discrepancy in our experimental results between vertical separation and the hybrid organizational structure, as predicted by model-theory, is surprising and can largely be attributed to the relatively high investments in the separated model. However, investment in the separation case exceeded that of the hybrid case by roughly 1.05 (p = .002). This result was driven partly by the final three rounds. A detailed examination of total investment indicates that in 9 of 12 rounds, average total investments do not differ significantly from each other. Figure 6 further reveals a parallel development of total investment in the hybrid and separation cases in rounds 1 to 7. Whereas, from round 8 onwards, investment in the separation case even increases, investment in the hybrid case decreases simultaneously.[17]

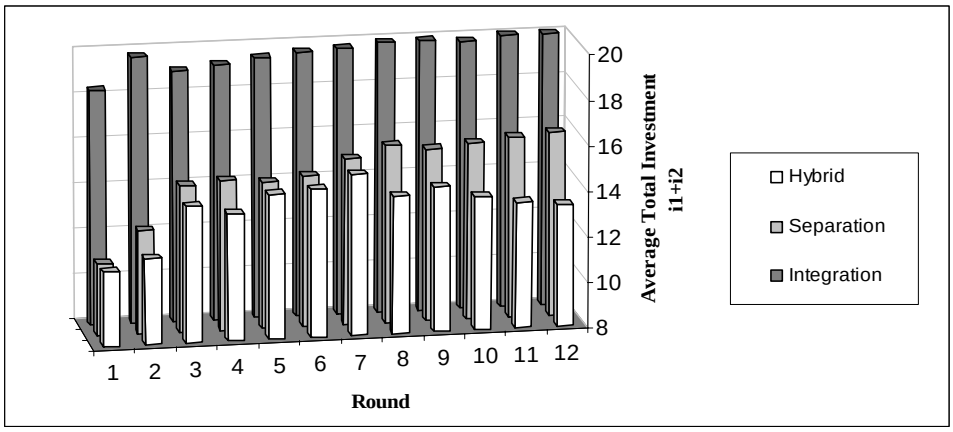

Figure 6. *Average Total Investments*

In the single-period treatments, 40 subjects participated in the separated case, 44 in the hybrid case and 42 in the integration case. An examination of investments in the single-period treatments yields similar findings to those reported for the multi-period treatments (see Figure 7). Although the results seem to match the rank order of theoretical predictions more closely, in-

[17] This growing discrepancy might originate from some form of last-period effect, as the test participants were asked to take part in at least 8 rounds. Last-period effects result in uncooperative player behavior in the final rounds of repeated games, because misbehavior cannot be sanctioned. However, since subjects were randomly matched with other partners in each round, direct sanctioning was not feasible anyway. Nonetheless, F_1s might have used underinvestment as a collective sanctioning device, this in fact losing credibility with an increasing probability of termination.

vestments do not differ significantly between the separation case (12.75) and the hybrid case (14.32). Average investments in rolling stock and rail infrastructure of the integration case (19.71) remain significantly superior and close to the social optimum. Yet, it is noticeable that investments in the first round of the multi-period games – 11.27 in the hybrid case, 11.18 in the separated case, and 18.33 in the integrated case – do not reach comparable levels. Thus, to some extent, the subjects may have relied on learning-by-doing in the multi-period treatments; this in turn indicates some kind of randomness in early-round decisions.

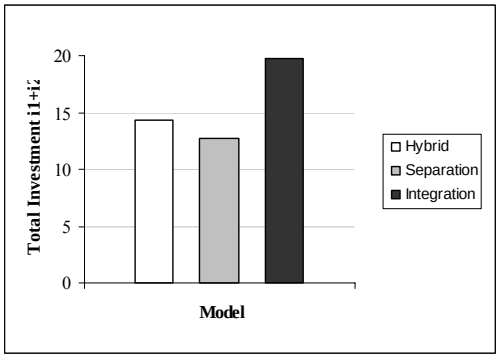

Figure 7. *Total Investments, Single-Period Treatments*

7. Discussion and Further Results

The results of our experimental analysis largely confirm the theoretical predictions of investment behavior in different institutional arrangements in the railway sector. In a world of incomplete contracts and asset specificity, with respect to specific investments, full vertical integration is the superior organizational solution. The amounts invested by our respondents were closer to the social optimal values than the amounts invested in a hybrid or separated structure. However, while investments in the integration scenario almost reach the predicted value and investments in the hybrid scenario differ by 23 percent, investments in the separation case substantially exceed the predicted levels.

The existence of social preferences may explain this apparent contradiction. In contrast to the utility functions of rational and self-interested actors, the utility functions of actors exhibiting social preferences also comprise the utility of the exchange partner. One potential out-come might be that actors do not consider investments as sunk at the time of negotiation, but expect the net benefits from investment to reflect the contribution of each player to the gross surplus, that is to investment cost (Homans, 1961; Selten, 1978).

	OLS	IRLS Robust Regression
$i_1/(i_1+i_2)$	0.566*** (0.071)	0.777*** (0.001)
Constant	0.205*** (0.356)	0.110*** (0.005)
N	264	264
F(1, 262)	63.880***	8954.460***
R^2	0.439	0.439

Dependent variable: F_1's share of surplus *b*. Robust standard errors in parentheses. Significance levels: *** p < 0.001; ** p < 0.01.

Table 3. *Estimation Results F_1, Separated Model*

An investigation of the bargaining outcomes indicates that, in fact, a significant proportion of respondents do index the bargaining behavior to prior investments. In order to test whether equity theory can contribute to the explanation of observed investment behavior, we estimate the

following simple equation: with F_1's share b as the dependent variable and his relative contribution to the total surplus as the independent variable. Table 3 displays the results of our estimations. We also applied a robust estimation technique to test the robustness of the estimation (Hackett, 1993; Hamilton, 1991). Both estimations strongly support the notion that the individual contribution to gross surplus exerts a substantial influence on the outcome of bargaining. Hence, social preferences seem to influence investment incentives and might explain the observation of higher investment levels than those predicted theoretically.

A further result of our experiment suggests that equal power due to shared ownership, as in the separation model, leads to efficiency losses, because of negotiation costs. In the sub-game perfect equilibrium, negotiations are successful and immediate, with bargaining terminating after the first round with breakdowns generally not being observed. This conclusion is based on the assumption of a homogenous group of actors. It follows that, since the sample becomes more heterogeneous, this result can no longer be retained, since rational actors might be disciplined by those with social preferences. Specifically, it may become rational to deviate from the theoretical prediction when confronted with potential and unexpected negotiation breakdowns by fair actors. Accordingly, in 24.6 per cent of cases, bargaining proceeds beyond round one. This causes efficiency losses of 25.7 per cent. In particular, total investments in all rounds amounted to 3.881, which correspond to a net joint profit of 19.405. A profit of only 14.417 could yet be realized, due to agreement delays and negotiation break-downs (Joskow, 2003).

8. Conclusions

The present chapter constitutes an initial contribution to the empirical analysis of organizational structures and investment incentives in the railway sector. In general, our experimental results seem to indicate that, in a world of incomplete contracts, a vertical separation of railways as well as hybrid forms might cause deficits in innovation, quality and safety, due to underinvestment in relation-specific assets. Although the levels of investments in the separation and the hybrid scenario exceeded the theoretical predictions, they failed to match the social optimum.

One of the main objectives of the European railways is the re-vitalization of rail traffic, which involves strengthening railway competitiveness in an intermodal comparison. The system's high technical and organizational complexity impedes or at least limits the potential for complete contracts which cover every conceivable aspect of the transaction. In order to determine which organizational structure is most appropriate for macroeconomic purposes, it is of paramount importance to consider the numerous effects of these various structures. Not only are incentives to invest relevant, but also aspects such as additional synergy effects, economies of scope, competition, subsidies, privatization revenues or the marketability of the railway industry. However, the problems that occurred in Great Britain after the railway restructuring process indicate that, particularly in this sector, considering the incentives to invest in innovation, quality and safety may be a particularly important aspect, and one that has so far been under-rated.

Our analysis shows that individual profit-maximizing behavior may lead to suboptimal decisions from a macroeconomic perspective, which in turn could constrain the competitive capabilities of European railways. Our results are also of particular importance for the design of hybrid models. The case considered here involves the separation of the right to decide on investments and the right to control the respective asset within the production process. From this, it follows that the investing party did not have any sanctioning potential after the investment had been undertaken, which may explain the comparatively low investments in rolling stock. However, this constellation constitutes only one possible alternative to designing a hybrid organizational form. Consequently, other design options, which could potentially combine the advantages of

the pure organizational forms more effectively, should be taken into account. In order to determine definitively which organizational structure is macro-economically superior, further quantitative-oriented research is required, especially to investigate the industrial-economic effects. However, it is advisable firstly to analyze the full impact of different organizational structures before making policy decisions.

9. References

Affuso, L. and Newbery, D. (2002). Investment, Reprocurement and Franchise Contract Length in the British Railway Industry. Discussion Paper No. 2619, Centre for Economic Policy Research.

Alchian, A.A. and Demsetz, H. (1972). Production, Information Costs, and Economic Organization. *American Economic Review*, 63: 777-795.

Bitzan, J.D. (2003). Railroad Costs and Competition: The Implications of Introducing Competition to Railroad Networks. *Journal of Transport Economics and Policy*, 37: 201-225.

Bolton, G. and Ockenfels, A. (2000). A Theory of Equity, Reciprocity and Competition. *American Economic Review*, 90: 166-193.

Bouf, D., Crozet, Y. and Leveque, J. (2005). Vertical Separation, Disputes Resolution and Competition in Railway Industry. 9th International Conference on Competition and Ownership in Land Transport (Thredbo 9), Lisbon, September 2005.

Cantos, P. (2001). Vertical Relationships for the European Railway Industry. *Transport Policy*, 8: 77-83.

Cantos, P. and Campos, J. (2005). Recent Changes in the Global Rail Industry: Facing the Challenge of Increased Flexibility. *European Transport*, 29: 1-21.

Crocker, K.J. and Masten, S.E. (1991). Pretia Ex Machina? Prices and Process in Long-Term Contracts. *Journal of Law and Economics*, 34: 69-99.

Crozet, Y. (2004). European Railway Infrastructure: Towards a Conversion of Infrastructure Charging? *International Journal of Transport Management*, 2: 5-15.

Driessen, G., Lijesen, M. and Mulder, M. (2006). The Impact of Competition on Productive Efficiency in European Railways. CPB Discussion Paper, 71.

European Commission (1991). Council Directive 91/440/EEC of 29 July 1991 on the Development of the Community's Railways.

European Commission (1996). A Strategy for Revitalising the Community's Railways. White Paper. COM (96) 421 final, 30 July 1996.

European Commission (2001). Directive 2001/12/EC of the European Parliament and of the Council of 26 February (2001) Amending Council Directive 91/440/EEC on the Development of the Community's Railways.

Fehr, E. and Fischbacher, U. (2002). Why Social Preferences Matter – The Impact of Non-Selfish Motives on Competition, Cooperation and Incentives. *The Economic Journal*, 112, C1-C33.

Ferreira, L. (1997). Rail Track Infrastructure Ownership: Investment and Operational Issues. *Transportation*, 24: 183-200.

Friebel, G., Ivaldi, M. and Vibes, C. (2003) Railway (De)Regulation: A European Efficiency Comparison. IDEI Report No. 3 on Passenger Rail Transport, University of Toulouse.

Gomez-Ibanez, J. (2004). Railway Reforms: An Overview of the Options. Conference on Railway Reform Rafael del Pino Foundation Madrid, 2004.

Grossman S. and Hart, O. (1986). The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration. *Journal of Political Economy*, 94: 691-719.

Growitsch, C. and Wetzel, H. (2009). Testing for Economics of Scope in European Railways: An Efficiency Analysis. *Journal of Transport Economics and Policy*, 43: 1-24.

Hackett, S.C. (1993). Incomplete Contracting: A Laboratory Experimental Analysis. *Economic Inquiry*, 31: 274-297.

Hamilton, L.C. (1991). *Regression with Graphics: A Second Course in Applied Statistics.* Pacific Grove, CA: Brooks/Cole.

Hart, O., (1995). *Firms, Contracts and Financial Structure.* New York, NY: Oxford University Press.

Hart O. and Moore, J. (1988). Incomplete Contracts and Renegotiation. *Econometrica*, 56: 755-785.

Hart, O. and Moore, J. (1990). Property Rights and the Nature of the Firm. *Journal of Political Economy*, 98: 1119-1158.

Hirschhausen C. von and Siegmann, J. (2004). SYNETRA: Synergien zwischen Bahnnetz- und -transport, Praxis, Probleme, Potentiale. Survey on behalf of the Bundesministerium für Bildung und Forschung, Berlin.

Homans, G.C. (1961). *Social Behavior: Its Elementary Forms.* New York: Harcourt, Brace and World.

International Union of Railways (UIC) (2003). Implementing the European Train Control System ETCS – Opportunities for European Rail Corridors. UIC ERTMS Conference, Leipzig, 10-11 December 2003.

Joskow, P.L. (2003). Vertical Integration. In Menard, C. and Shirley, M.M. (Eds.) *Handbook of New Institutional Economics*. Dordrecht: Springer.

Klein, B. Crawford, R. and Alchian, A., (1978). Vertical Integration, Appropriable Rents, and the Competitive Contracting Process. *Journal of Law and Economics*, 21: 297-326.

Link, H. (2003). Rail Infrastructure Charging and On-Track Competition in Germany – Nine Years Later. European Transport Conference, Strasbourg, 2003.

Maskin, E.S. and Moore, J. (1999). Implementation and Renegotiation. *Review of Economic Studies*, 66: 39-56.

Merkert, R., Nash, C.A. and Smith, A.S.J. (2008). Looking Beyond Separation – A Comparative Analysis of British, German and Swedish Railways from a New Institutional Perspective. European Transport Conference. Leiden, 2008.

Nash, C.A. (2005). Rail Infrastructure Charges in Europe. *Journal of Transport Economics and Policy*, 39: 259-278.

Nash, C.A. (2006). Europe: Alternative Models for Restructuring. In Gomez-Ibanez, J.A. and de Rus, G. (Eds.) *Competition in the Railway Industry: An International Comparative Analysis.* Northampton, MA: Edward Elgar.

Nash, C.A. and Preston, J.M. (1994). Competition in Rail Transport: A New Opportunity for Railways. Working Paper 397, Institute for Transport Studies, University of Leeds.

Nash, C.A. and Rivera-Trujillo, C. (2007). Rail Reform in Europe: Issues and Research Needs. In Rietveld, P. and Stough, R.R. (Eds.) *Institutions and Sustainable Transport: Regulatory*

Reforms in Advanced Economies. Cheltenham, UK and Northampton, MA, USA: Edward Elgar.

Nash, J.F. (1950). The Bargaining Problem. *Econometrica*, 18: 155-162.

Nöldeke G. and Schmidt, K.M. (1998). Sequential Investments and Options to Own. *RAND Journal of Economics*, 29: 633-653.

Pfund, C. (2003). The Separation of Infrastructure and Operations Constitutes a Fundamental Mistake. *Public Transport International*, 52: 32-33.

Pittman, R.W. (2005). Structural Separation to Create Competition? The Case of Freight Railways. *Review of Network Economics*, 4: 181-196.

Pittman, R.W. (2007). Make or Buy on the Russian Railway? Coase, Williamson, and Tsar Nicholas II. *Economic Change and Restructuring*, 40: 207-221.

Preston, K. (1996). The Economics of British Rail Privatisation: An Assessment. *Transport Reviews*, 16: 1-21.

Rabin, M. (1993). Incorporating Fairness into Game Theory and Economics. *American Economic Review*, 83: 1281-1302.

Roth, A.E. (1995). Introduction to Experimental Economics. In Kagel, J.H. and Roth, A.E. (Eds.) *Handbook of Experimental Economics*. Princeton, NJ: University Press Princeton.

Rothengatter, W. (2001). How Good is First Best? Marginal Cost and other Pricing Principles for User Charging in Transport. Working Paper, University of Karlsruhe.

Rubinstein, A. (1982). Perfect Equilibrium in a Bargaining Model. *Econometrica*, 50: 97-109.

Rus, G. de and Nombela, G. (2007). Is Investment in High Speed Rail Socially Profitable? *Journal of Transport Economics and Policy*, 41: 3-23.

Sanchez, P.C., Monsalvez, J.M.P. and Martinez, L.S. (2008). Vertical and Horizontal Separation in the European Railway Sector: Effects on Productivity. Documentos de Trabajo, Fundacion BBVA. 12.

Selten, R. (1978). The Equity Principle in Economic Behaviour. In Gottinger, H. and Leinfellner, W. (Eds.) *Decision Theory and Social Ethics, Issues in Social Choice*. Dordrecht: Reidel.

Sloof, R. (2004). Finite Horizon Bargaining with Outside Options and Threat Points. *Theory and Decision*, 57: 109-142.

Smirnov, V. and Wait, A. (2004). Hold-Up and Sequential Specific Investments. *RAND Journal of Economics*, 35: 386-400.

Williamson, O.E. (1975). *Markets and Hierarchies: Analysis and Antitrust Implications.* New York: Free Press.

Williamson, O.E., (1985). The Economic Institutions of Capitalism. New York: Free Press.

Yvrande-Billon, A. (2004). Franchising Public Services: An Analysis of the Duration of Passenger Rail Franchises in Great Britain. In Cliquet, G., Hendrickse, G., Tuunanen, M. and Windsperger, J. (Eds.) *Economics and Management of Franchising Networks.* Heidelberg: Physica.

II. THE COSTS AND BENEFITS OF NETWORKING*

1. Abstract

This chapter aims at extending the scope of the current discussion on networking activities by (1) providing a brief conceptualization of important elements of both networking benefits and costs and (2) considering how different networking activity levels may affect optimal institutional choice, taking possible non-linear developments and opportunity costs into account. By doing so, one of the main intentions of this chapter is to point to potential downsides of networking – an issue that has largely been neglected so far.

* This chapter originates from joint work with Thomas Ehrmann. An early version of this chapter was presented at the EMNet 2007 Conference in Rotterdam, NL. See: Ehrmann, T. and Schmale, H. (2007), The Costs and Benefits of Networking. Working Paper presented at the International Conference on Economics and Management of Networks, 2007, Rotterdam, NL.

2. Introduction

"Your true value depends entirely on what you are compared with."
Bob Wells

A fast growing body of literature, emerging from a wide array of academic perspectives, deals with the subject of hybrid governance modes, in particular, with network forms of organization (for an overview, see Academy of Management Review, Special Issue 2006). This literature intends to shed light on the question whether to produce a certain good or service in-house or not, thereby addressing a fundamental managerial challenge, i.e., the choice of organizational structure. Yet, we feel that one important shortcoming characterizes large parts of the ongoing debate on network forms of organization: dysfunctionalities and costs of networking truly remain underexplored. In this study, we aim at extending the scope of the current discussion on networking activities by (1) providing a brief conceptualization of important elements of both networking benefits and costs and (2) considering how different networking activity levels may affect optimal organizational choice, taking possible non-linear developments in benefits and costs into account. By doing so, one of our main intentions is to point to potential downsides of networking – an issue that has largely been neglected so far.

Previous studies appear to take an overly optimistic perspective on hybrids, mostly focusing on beneficial effects of networking (Gargiulo and Benassi, 1999; Labianca and Brass, 2006; Pittaway et al., 2004). Organizational scholars and sociologists, for instance, often certify network forms of organization superior functionality regarding knowledge transfer, resource acquisition, and legitimization issues (e.g., Baum and Oliver, 1991; Gulati, 1995; Pillai, 2006; Powell, Koput and Smith-Doerr, 1996). Yet, ignoring dysfunctionalities and costs of networking hinders an in-depth understanding of how collective action could be organized effectively (Parkhe, Wasserman and Ralston, 2006). In the same vein, Podolny and Page (1998: 66) have observed that the "attention to the functionality of network forms of organization explains why economic actors rely on network forms of organization, but it does not explain why they do not." Moreover, they argue that scholars' widespread approach to concentrate on analyzing the benefits of networking might be useful for explaining varying reliance on networks *between* industries,

since the relevance of collaboration-related benefits, such as learning and knowledge transfer, could differ from sector to sector. However, it still remains unclear why a considerable fraction of actors *within* an industry may constantly rely on hierarchical governance modes without investing in network building.

Although not questioning that networking entails considerable costs (e.g. Burt, 2005), scholars have mostly refrained from elaborating on a conceptual or formal framework which incorporates *both* the benefits and the costs of networking.[19] As a more recent example, Lavie (2006) has developed a theoretical model which deals with the formalization of competitive advantages derived from access to external resources, i.e., network relations. In extending the resource-based view of the firm, Lavie contributes to further our understanding of how economic rents are generated by applying resources which originate from outside the firm. However, similar to other common approaches to understanding networking activities, two important restrictions limit the testability as well as the explanatory power of Lavie's model. First, Lavie implicitly assumes a linear-additive relationship between additional shared resources, i.e., a rising number of network partners, and a focal firm's extracted rent. Despite that the author acknowledges that network relations might vary in performance, he missed the opportunity to explicitly incorporate this fact into his model. This is surprising since recent studies show evidence on *non-linear* relationships between a firm's network activity level, i.e. the number of network partners, and the benefits obtained from external resources (Uzzi, 1996; 1997). The second restriction concerning the Lavie-model is the lack of an explicit cost consideration (Lavie, 2006: 651). Neglecting costs and dysfunctionalities of networking breeds the assumption that networking always secures positive rents. Consequently, the following major question emerges: *Why do not all economic actors perceive networks as a panacea and deliberately choose other organizing forms from the outset?*

[19] The transaction cost approach has been criticized for not explicitly comparing the costs *and* benefits of institutional choice (e.g. Zajac and Olsen, 1993), although Riordan and Williamson (1985) in fact do integrate transaction as well as production cost differences within their model. Since transaction costs are a key aspect of institutional efficiency in network arrangements, we base significant parts of our argumentation on ideas from TCE.

Based on insights from management and social network theory literature, we here take an initial step in addressing the questions to which extent and under which circumstances the network form of organization represents a superior organizational choice. In a nutshell, we first emphasize that considering *both* networking benefits and costs is crucial to substantiating the economic calculus which underlies a comparative analysis of governance modes. As we will outline in detail in the next section, important sources of benefits are knowledge transfer, relation specific investments (e.g. in dedicated machinery or human capital), and access to complementary and/or scarce resources (e.g. distribution network or reputation). Important costs involved in networking are transaction costs, originating from the need to control potentially opportunistic partners (e.g., White and Siu-Yun Lui, 2005; Williamson, 1996) and coordination costs related to communication and coordination efforts independent from incentive conflicts (Camerer and Knez, 1997; Gulati and Singh, 1998; Park and Ungson, 2001).

Second, and in line with findings from recent research on networking, we make a strong case for non-linearities in the relationship between the degree of collaborative activity and the costs and benefits of networking. By doing so, we explicitly account for varying value effects of different activity levels within the governance mode of networking since intensity of usage has been revealed as a crucial determinant of optimal institutional choice (Burt, 2005; Uzzi, 1997).

Third, and finally, the consideration of positive opportunity costs – being defined as foregone net benefits from the most attractive alternative resource employment (e.g. Varian, 1997) – is a necessary precondition for quantifying the competitive advantage a particular governance form actually offers. The disregard of an opportunity cost calculus would impede the comparison of alternatively available governance modes, and consequently, the identification of a true competitive advantage.

The remainder of this chapter is organized as follows. In the next section we outline the general economic rationale of the choice of organizational structure. Then in section 4, we turn to the sources of networking benefits and furthermore explain how different activity levels can affect the potential benefits derived from external relations. Section 5 deals with the sources and non-

linear developments of networking costs. Moreover, we introduce an opportunity cost calculus. In section 6, a simple heuristic model summarizes the ideas and illustrates that too extensive a reliance on network resources could raise the attractiveness of the opportunity, implying a competitive disadvantage for interconnected firms. Section 7 concludes.

3. Weighing the Benefits and Costs of Networking

Scholars emphasize that data on alliance failure is hard to obtain, yet, the scarce empirical evidence reports on a very high percentage of interfirm collaborations (more than 50%) which either do not operate in the way they are supposed to or even fail (Podolny and Page, 1998; Park and Ungson, 2001; Parkhe, 1993). Hence, one should assume that in some cases alternative governance modes have become superior relative to the network form in terms of costs and benefits, implying a competitive *disadvantage* for interconnected firms.

Here, we define a network form of organization as any collection of actors ($n \geq 2$) that pursue repeated enduring exchange relations with one another and, at the same time, lack a legitimate authority to arbitrate and resolve disputes between the actors. In this respect, the network form of organization constitutes a discrete organizational form next to markets and hierarchies, with distinctive adaptation and control mechanisms (Podolny and Page, 1998; Powell, 1990).

In particular, trust, i.e., an actor's confidence that another actor will not exploit its vulnerabilities (Sako, 1992), and reciprocity, i.e., a conditional behavioral pattern that returns ill for ill as well as good for good (Axelrod, 1984), are commonly alleged to play decisive roles as governance mechanisms especially in the network form of organization (Powell, 1990). Moreover, networking lays the foundations for the development of social norms. A social norm exists when the socially defined right to control an action is held not by the actor but by others (Coleman, 1990). Repeated interaction between densely interconnected partners breeds a collective understanding of generalized norms of cooperation such as fairness, team-work, or openness, for example. Given the danger of collective sanctioning in the case of defection, social norms can provide strong incentives for cooperation.

From any utilization of different means of governance follows both a difference in cost and benefit generation. Hence, an analysis of governance forms that concentrates on either benefits or costs oversimplifies. In pursuing this notion, Windsperger (1987), for instance, called for a combined analysis of costs and benefits from the utilization of different governance modes referring to the case of franchising where institutional arrangements are sometimes explicitly cho-

sen for revenue maximization reasons rather than cost-economizing objectives. Thus, for a governance form to be superior it must be cost-efficient in terms of adapting, coordinating, and safeguarding exchanges (e.g. Williamson, 1991), taking differences in benefit generation into account (Dyer, 1997; Ring and Van de Ven, 1992; Zajac and Olsen, 1993).

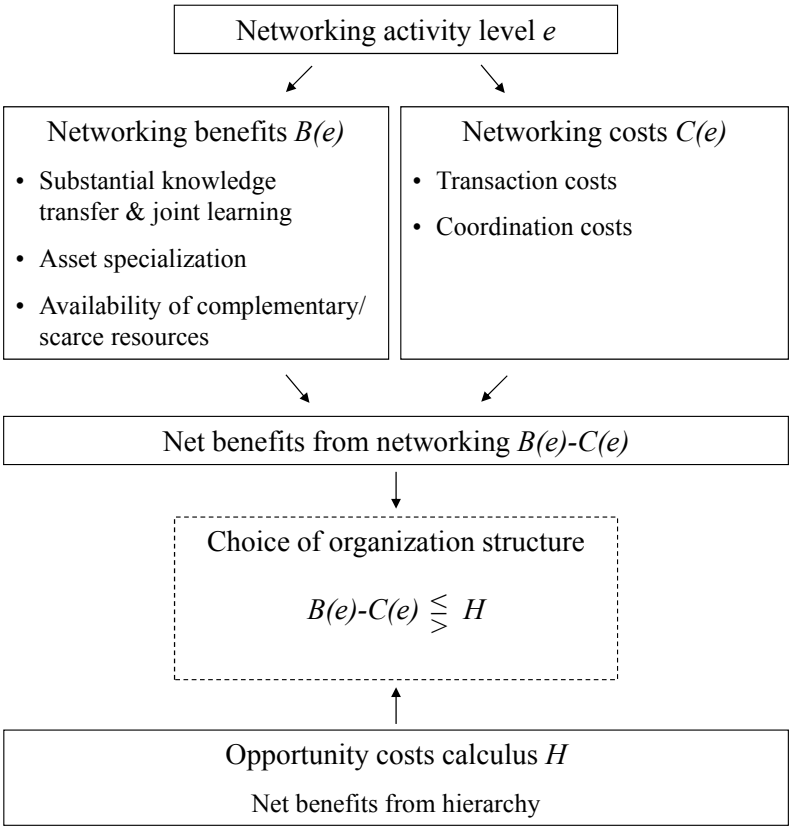

Figure 8. *Decision Framework of Organizational Choice*

Basically, the derivation of a decision rule for the choice of an adequate organization structure is straight forward. For instance, if $B(e)$ denotes the benefits and $C(e)$ the costs of networking at an activity level *e,* and *H* the opportunity costs (i.e., net benefits from the next best alternative governance mode, see Ch. 5.3), then the focal firm should opt for networking if $B(e) - C(e) > H$ (see Figure 8). In other words, as long as the benefits associated with choosing the network form exceed the *entire* costs one incurs by doing so, this is an advantageous choice (Frank, 2000). However, as simple as this decision rule appears, as complex the application of this rule can get, since in the first place it requires an identification and estimation of the involved benefits and (opportunity) costs.

4. The Benefits of Networking

4.1 Sources of Networking Benefits

Studying cooperation-related benefits, scholars have highlighted the relevance of aspects such as legitimization, learning, risk reduction or the combination of complementary skills (e.g., Ahuja, 2000; Baum and Oliver, 1991; Eisenhardt and Schoonhoven, 1996; Powell, Koput and Smith-Doerr, 1996). In an attempt to structure the variety of findings, Dyer and Singh (1998) established an integrated framework which explains how firms capitalize on collaboration activities. In particular, important benefits from networking can be subsumed under the following broader categories: (1) knowledge exchange and joint learning, (2) investments in relationship-specific assets and (3) availability of complementary and/or scarce resources. These benefits translate into profits either due to production cost-minimizing effects or due to the maximization of revenues (Dyer and Singh, 1998).

Perhaps the most emphasized reason for interfirm collaboration is the feasibility of knowledge transfer and joint learning (e.g. Powell, Koput and Smith-Doerr, 1996; Uzzi, 1997). Albeit external relationships are primarily perceived as a key vehicle for information transfer leading to substantial product innovations (e.g., sales partner feedback), they can also be critical for promoting awareness and early adoption of new business practices developed by other organizations, which might lead to enhanced efficiency (Pittaway et al., 2004: 145). In particular, enduring exchange relations between partners create room for the development of routines allowing for the transmission of tacit knowledge, i.e., knowledge which cannot easily be codified and transferred via arm's-length relationships (Kogut and Zander, 1992). Tacit knowledge, that component of knowledge which defines the knowing-how rather than knowing-what, is of particular value since it is difficult to imitate and therefore provides the potential for competitive advantages (Dyer and Singh, 1998). Furthermore, information transfer can initiate learning

processes which improve the involved firms' skills to respond to environmental changes, thereby enhancing adaptation capacities in the face of uncertainty (Pillai, 2006: 132).[20]

Both the long-term character of network relations and the distinctive means of governance build important foundations for the willingness to pursue partner-specific asset specialization, a widely recognized source of rent generation (e.g. Amit and Schoemaker, 1993). Benefits from investments in relationship-specific assets arise from substantial productivity gains within the entire production process (Williamson, 1985). For instance, physical asset specialization, such as the investment in dedicated machinery, can provide potentials for product quality improvements as well as an increase in output due to a reduction of production steps or declining failure rates. Both the specializations of human assets, e.g., investments in the accumulation of partner-specific know-how, and investments in site-specific assets, i.e., the establishment of physical proximity, provide quality enhancement and economizing potentials in a similar way (Dyer, 1997).

Finally, another widely accepted source of networking benefits is the availability of complementary and scarce resources, complementary resources being defined as a set of resources which collectively generate higher benefits than the sum of benefits obtained from their individual use (Hamel, 1991; Dyer and Singh, 1998). Thereby, the combined resource endowments are more valuable, rare, and difficult to imitate (Barney, 1991) than they had been before they were combined (Dyer and Singh, 1998). Some kinds of resources cannot be acquired via market transactions or cannot be developed by firms themselves in the short run. For instance, reputation (Baum and Oliver, 1991), specialized expertise (e.g. access to product development skills), or other tacit skills are slow to develop in-house (Eisenhardt and Schoonhoven, 1996; Hennart, 1991) and not necessarily available on markets because of indivisibility (Oliver, 1997). Since differential historical patterns of firm evolution and path dependencies promote differences in organizations' asset endowments (e.g. Alchian, 1950), needed resources might nevertheless

[20] In this context, Uzzi emphasizes fine-grained information transfer as a main advantage of embedded relations. "Information exchange in embedded ties is more proprietary and more tacit than the information exchanged at arm's length. It includes strategic and tacit know how that boosts a firm's transactional efficacy and responsiveness to the environment." (Uzzi, 1996: 678).

exist within the network. Hence, interfirm relationships can serve as conduits to provide for difficult-to-acquire and complementary assets which provide potentials to enhance a firm's economic performance (Sakakibara, 1997).

As we will show in more detail below, the potential amount of benefits obtained from networking, i.e. from knowledge exchange, relation-specific investments, and complementary and/or scarce resources should crucially depend on the level of networking activity. Possible non-linear developments in benefit generation will be discussed in the next section.

4.2 Non-Linear Development of Networking Benefits

Numerous studies implicitly assume a monotonically increasing relationship between firms' degrees of interconnectedness and the value effects of additional shared resources. We have concerns about this view since several arguments indicate that this assumption may not hold for the full range of network activity levels.

We argue that, *at low levels* of activity, the potential amount of benefits obtained from networking should be high and potentially increasing in network activity. Generally, the acquisition of complementary or scarce resources such as reputation or specialized expertise is valuable to the firm. The benefits are supposed to be high at the outset because of a relatively high probability that the focal firm gains access to new information and needful resources with every new connection. However, central to taking advantage of a munificent networking environment is the capability to locate and to recognize complementary and scarce resources. Growing collaboration experience supports the formation of evaluation and environmental screening capabilities (Dyer and Singh, 1998). Moreover, third-party interaction can enhance the understanding of partner-specific capabilities, thereby breeding a collective knowledge of network members' skill endowments. Both the improvement of evaluation and screening capabilities and decreasing information asymmetries within the network facilitate the identification of complementary resources and information, potentially leading to a more effective use of network ties and increasing benefits (Argote, McEvily and Reagans, 2003).

In a similar vein, Cohen and Levinthal (1990) emphasize that the ability to identify, absorb, and exploit external knowledge, i.e., the amount of absorptive capacity, depends on the prior possession of issue-related knowledge. Interaction with an existing base of partners can enhance the accumulation of issue-related knowledge, and consequently, improved absorptive capacity can lead to greater effectiveness in knowledge internalization and utilization (Powell, Koput and Smith-Doerr, 1996).

Potential gains from relation-specific investments might also rise at the outset. This stems from the fact that increasing interaction and a growing number of partners should initially support the development of trust and social norms as informal governance mechanisms that safeguard exchange in networks.[21] Incentives to invest in partner-specific assets generally are tempered by the omnipresent danger of exploitation by opportunistic partners (Klein, Crawford and Alchian, 1978). For instance, exploitation might take the form of re-negotiation of pay-off allocation agreements after investments are sunk. Hence, underinvestment is the dominant strategy as long as there are no sufficient safeguards in place (e.g. Hart, 1995). But when informal safeguards become more functional, eventually filling the gaps of necessarily incomplete formal contracts, the incentives to underinvest are reduced.[22]

However, *too high levels* of interconnectedness and a too high number of partners might eventually lead to reverse effects, implying a high probability of declining benefits after a certain threshold of activity levels.

First, with respect to the access to complementary and scarce resources, higher numbers of network partners may increase the probability that the firm obtains irrelevant or even redundant resources (Witt, 2004).[23] The same applies for benefits from knowledge exchange: a too large

[21] Trust and social norms are built up not only by direct interaction but also via the assessment of potential partners' behaviors within third-party interactions (Henrich, 2004; Henrich and Henrich, 2006).
[22] Hart (1995) emphasizes that formal contracts are necessarily incomplete. First, bounded rationality and prohibitively high transaction costs prevent from the formulation of complete contracts. Second, information asymmetries limit external enforcement of agreements.
[23] Moreover, Podolny and Page (1998) point to the fact that firms are limited in their ability of adapting network configuration due to constraints on the dissolution of ties because of reputational costs. As long as the ability to from ties with new partners is contingent on the number of preexisting ties (Gulati, 1995), a tie which is no longer beneficial might persist for the purpose of signaling good reputation.

base of network ties comes along with an increasing grade of information redundancy and inefficiencies in information flows due to overembeddedness, i.e. a too high number of network relations, may arise (Uzzi, 1996; 1997).

Second, a too high number of network partners may hinder an intensive exploitation of connections, which eventually leads to decreasing benefits from further network extensions. This follows from limited management capacities (Granovetter, 2005). With an increasing number of network partners, the focal firm may fail to maintain adequate (based on partners' expectations) levels of cooperation intensity. Cooperation intensity reflects several dimensions of an exchange such as the frequency of contact and the extent to which and what resources and information are transferred. Intense relations encourage the exchange of complex and tacit knowledge (Nahapiet and Ghoshal, 1998). Moreover, strong relations are reliable in terms of assistance and typically allow a better and more frequent recourse to external resources (Burt, 2005). Considering the equity rationale underlying reciprocity norms, i.e., an actor in an exchange relation will expect the rewards of each party to be proportional to his input (Homans, 1961), external partners are likely to respond in terms of low contributions likewise.

Third, the functionality of social norms and trust as means to safeguard specific investments are likely to decrease after a certain threshold of activity levels. In particular, increasing network size might lead to the breakdown of informal institutions (Davis, 2006). Granovetter (2005: 34) states accordingly that "the larger the group, the lower its ability to crystallize and enforce norms". This in turn might mitigate incentives to invest specifically.

5. The Costs of Networking

5.1 Sources of Networking Costs

Networking costs result from the establishment, maintenance, exploitation, and abandonment of exchange relations. Important costs can be subsumed under the broader categories of (1) transaction costs, originating from the need to control potentially opportunistic partners and (2) coordination costs, related to communication and coordination efforts independent from incentive conflicts. Moreover, our argumentation puts a strong emphasis on the consideration of opportunity costs which form another source of networking costs. By reason of its particular importance, the issue of opportunity costs has been dedicated a chapter of its own (Ch. 5.3).

In order to prevent negative effects due to opportunistic appropriation by external partners, firms incur considerable transaction costs in stipulating various formal and informal agreements (Williamson, 1985). For instance, partners might tempt to unilaterally reduce efforts in joint operations, distort or withhold information, or defect from other kinds of mutual agreements at the other firm's expense (Hart, 1995; Lavie, 2006; Parkhe, 1993). Ex-ante transaction costs occur during the initiation process of a transaction, e.g., costs of acquiring information on potential partners, as well as within the contracting process, e.g., costs of negotiating and safeguarding an agreement. Ex-post transaction costs stem from monitoring efforts, from potential haggling and from ex-post adaptations and re-alignments of agreements due to environmental changes (Williamson, 1985). Scholars have argued that network organizations' reliance on trust and commonly shared norms as informal safeguards can translate into significantly lower transaction costs as compared to governance modes primarily relying on formal safeguards, such as legal contracts (e.g. Dore, 1983; Sako, 1992; Uzzi, 1997). For instance, trust may both lower ex-ante and ex-post transaction costs by facilitating agreement processes and by lowering monitoring necessities, respectively (Dyer and Singh, 1998). In this vein, Jarillo has argued that "being able to generate trust, therefore, is the fundamental entrepreneurial skill to lower those costs and make the existence of the network economically feasible" (Jarillo, 1988: 36). However, building informal safeguards such as trust and commonly shared norms is costly for itself. It requires a

considerable up-front investment of at least time, and trust also needs to be cultivated as the relationship matures (Dyer, 1997). The amount of investment into trust-building crucially depends on the parties' willingness to cooperate and on incentives to defect. In this context, scholars have emphasized the role of network structure as an important determinant of the institutional environment (Burt, 1992; Burt, 2005; Granovetter, 2005; Lorenzen, 2002; Pillai, 2006).[24] Small cohesive networks with low information transfer costs, for instance, can foster the emergence of trust given that reputation leverages group-conform behavior (Henrich, 2004; Williamson, 1996). Here, parties will try to protect their reputation by behaving cooperatively (Burt, 2005).[25]

Coordination costs refer to costs in terms of time, energy or money devoted to both communication and coordination between autonomous cooperation partners. Even in the absence of any opportunism, failures of adaptation may arise because autonomous parties could read external signals differently and choose different reactions to these signals, although the parties' common purpose might be to achieve timely and harmonized responses (Gulati, Lawrence and Puranam, 2005; Williamson, 1991). Further intensifying the need for alignment between the network members, differences in historical patterns and firm evolution are an important reason for differential resource endowments, capabilities, and organizational structures. Parties to a transaction might face significant communication and coordination efforts intending to overcome these organizational incompatibilities (Camerer and Knez, 1997; Park and Ungson, 2001; White and Siu-Yun Lui, 2005). Moreover, the more fine-grained specialization gets, the more complex and time consuming becomes the task to coordinate individual efforts. Hence, coordination costs are primarily driven by interpartner diversity and managerial complexity.

[24] Although being frequently criticized for under-emphasizing the role of social norms or trust as governance mechanisms, this is what Williamson refers to as the institutional environment, acknowledging that "in effect, institutional environments that provide general purpose safeguards relieve the need for added transaction-specific supports" (Williamson, 1996: 267). In a similar vein, Pittaway et al. state that "inclinations towards trust, opportunism, legal contracting and self-interest are all shaped by the institutional context in which firms operate" (Pittaway et al., 2004: 157).

[25] However, it is important to notice that the institutional environment is not static but rather evolving, subject to network structure and hence networking activity.

As we will show in detail below, the amount of costs incurred through networking should depend on the networking activity level. Possible, non-linear developments of networking costs are described in the next section before we will turn to the topic of opportunity costs thereafter.

5.2 Non-Linear Development of Networking Costs

At *low levels of networking activity* the costs of networking are supposed to be considerably high. A firm that initiates networking activities is likely to have little experience in formal contract negotiation and cannot resort to standardized action patterns in terms of managing external relations. Hence, setting up ties that do not have any precedent is very expensive (Burt, 2005). Networking firms have to bear costs originating from both the need to control transactions with potentially opportunistic exchange partners who might be trying to pursue self-interested objectives and the need to coordinate and align action of independent entities (White and Siu-Yun Lui, 2005).

In particular, firms need to gather a lot of information on the potential partner's line-up concerning objectives, resource base, cooperation abilities and reputation. Moreover, the omnipresent danger of opportunistic behavior implies high demands on transaction safeguards (Williamson, 1996). Hence, firms try to protect themselves by establishing causal ambiguity, registering trademarks, applying for patents or by building other costly isolating mechanisms. The need to coordinate action with external partners also implies high costs initially because the firm first of all has to establish communication routines, create organizational interfaces, and reorganize processes in order to enable joint action with external partners (White and Siu-Yun Lui, 2005). Relatedly, Dyer, Kale and Singh (2001) emphasize the importance of a dedicated management function as a key instrument for managing external relationships. However, dedicated functions imply considerable up-front investments and thus an organizational long-term commitment to affiliation strategies. Coordination costs in terms of time and effort devoted to direct interaction are also supposed to be high at the outset of any networking engagements. Differences in strategic goals (Doz, 1988), organizational cultures (Parkhe, 1991), and cognitive patterns have to

be identified and resolved by mutual adjustment, which incurs costs in terms of time and efforts devoted to communication and coordination (White and Siu-Yun Lui, 2005).

Initially, coordination and control costs may decline due to the development of efficient routines in coordinating partners, and experience effects in contract negotiation and monitoring, for example. Moreover, given indirect connections among partners, transaction cost reducing social mechanisms such as trust and commonly shared norms may be reinforced (Burt and Knez, 1995; Henrich and Henrich, 2006; Coleman, 1990; Gulati and Gargiulo, 1999; Lorenzen, 2002). Hence, within rather densely interconnected cliques, monitoring and enforcement efforts can to some extent be "sourced out" to the network, given an increasing likelihood that non-cooperative behavior is observed and sanctioned by common partners.[26] Consequently, collective monitoring and sanctioning, e.g., in the form of network exclusion, might reduce the focal firm's monitoring and enforcement costs (Rowley, 1997). Cheap third-party information on potential new partners made available by indirect ties might also lead to a reduction in costs of information acquisition (Granovetter, 1973).[27]

However, *at high levels of network activity* and with a growing number of links, it is realistic to assume that potential gains from experience are more than offset by increasing costs of coordination. First, increasing specialization raises the level of managerial complexity. Given that tasks bear connections to some extent, the focal firm incurs a high amount of costs in terms of time and effort to ensure the timely transfer of accurate information to each of the partners in case of adaptation, for instance. Second, a large number of partners raise the diversity of organizational configurations within the network, which inevitably leads to higher coordination costs when aligning actions. Consequently, after a certain threshold of network activity levels and with a growing number of links, the firm should experience higher coordination costs.

[26] The likelihood of dense interconnection within the network is maximized for small or moderate group sizes. Density is defined as the proportion of ties in a network relative to the total number possible (Wasserman and Faust, 1994). With growing network size, density decreases because actors are limited in management capacity and the number of possible ties increases exponentially (Granovetter, 2005).

[27] In a similar vein, Gulati (1995) found that the ability of network tie formation positively correlates with the number of preexisting ties. When searching for adequate affiliates, potential partners might be guided by other firms' histories. In particular, they tend to take the number of prior exchange relations as a reference for cooperative intentions.

A very large network might lead to the breakdown of informal institutions which eventually leads to rising costs of control (Davis, 2006). In particular, monitoring each partner's contribution, exchanging information on inadequate behavior, and imposing sanctions in the face of opportunistic conduct become increasingly time-consuming, since the risk of receiving inaccurate information about what potential partners did in the past raises with growing communities (Henrich, 2004). Hence, the lower the ability of the group to enforce norms (Granovetter, 2005), the higher may be the need for additional transaction specific safeguards.

5.3 Opportunity Costs

Any neglect of positive opportunity costs from alternative resource deployments would imply that the analysis of governance choice lacks a tertium comparationis. The consideration of opportunity costs is essential to assessing the true cost of any alternative. Conversely, ignoring opportunity costs may breed the false impression that its benefits are free of charge, thus leading to an erroneous evaluation of an alternative's actual performance.

In order to determine the competitive advantage of an interconnected firm, first one has to answer the basic question of what to compare to. Arend and Seale (2005) have emphasized the requirement of comparability of resource uses with respect to anticipated risk, financial and/or organizational commitment, and strategic objectives when defining an appropriate opportunity cost calculus. Hence, the authors propose different benchmarks as potential measures for opportunity costs. In case of a strategic alliance network, for example, the opportunity might be regarded as the net gains from an internal venture, i.e., the establishment of an organizational entity being subject to the ultimate authority of the firm and with access to all internal resources (Shortell and Zajac, 1988). Similarly, Contractor and Lorange have suggested anchoring the evaluation of collaborative efforts to an "internalization" or "go-it-alone" option (Contractor and Lorange, 1988) in order to reveal the benefits and costs endemic to cooperative relationships with external partners. Central to this logic is the fact that a firm always has the option to approach its strategic objectives independently and to dedicate its resources to internal domains. However, this choice could result in somewhat different outcomes, i.e., either higher or lower

returns, since networking and hierarchy are not complete substitutes to both the benefit and cost side (Conner, 1991; Contractor and Lorange, 1988; Williamson, 1985).

With respect to the cost side, for example, a firm that opts for internalization does not have to cope with opportunistic appropriation hazards and organizational and structural incompatibilities – the key cost drivers of networking activity. Certainly, hierarchy also causes transaction and coordination costs, yet the nature of these costs differs substantially. First, hierarchy relies on force majeure rather than formal contracting and trust-based collaboration as governance mechanisms (Williamson, 1985). Consequently, transaction costs do not vary in an activity induced development of commonly shared norms and trust. Second, the "go-it-alone" option does not involve the need to handle the wide spectrum of cultures, believes, and organizational structures which come along with an increasing need for costly communication and coordination efforts when synchronizing action within networks (Park and Ungson, 2001). Roughly the same logic applies for the benefit side, given the fact that internalization implies being "constrained" to internal assets or resources available through market acquisition. Consequently, representing a meaningful and rather simple benchmark, hierarchy may serve as a point of reference for evaluating the benefits and costs involved in networking and hence for determining an alliance network's superiority and related competitive advantages. In the following, we will summarize our ideas on the costs and benefits of networking in a simple heuristic model.

6. A Simple Heuristic Model

We assume both the costs and the benefits of networking to be a function of the effort devoted to distinct managerial network activities (White and Si-Yun Lui, 2005), the effort level being represented by the number of commitments (Lavie, 2006). The gradients and the position of the benefit and the cost curve are determined by the intermediation environment of the respective industry and are thus assumed exogenous. The intermediation environment affects the costs and benefits of networking and reflects characteristics such as information asymmetries, competitiveness, and uncertainty (Kali, 2003). Along with our discussion on the non-linear development of networking costs, the cost function of networks is defined as u-shaped: $C(e) = a \cdot (e - d)^2 + f$, with f, $d \geq 0$ and a, $e > 0$, where f denotes the lowest possible cost level corresponding to the activity level $e = d$. Accordingly, with regard to the gains of networking, we assume the benefit function to be inversely u-shaped $B(e) = -g \cdot (e - k)^2 + l$, with e, $g > 0$ and k, $l \geq 0$, where l denotes the maximum level of benefits from networking at the activity level $e = k$ (see Figure 9, left). We deliberately chose simple functions for substantiating the development of networking costs and benefits. Further differentiations, e.g., with regards to specific components of networking costs and benefits, are of course feasible but would not significantly add to the illustration of our arguments.

Based on this simple framework, the costs of networking are denoted by $C(e)$, and the benefits of networking are denoted by $B(e)$. The net benefits from networking $B(e) - C(e)$ can now be compared to the net benefits H from the use of hierarchy for organizing the same activities. Figure 9 (right) graphs the net benefit curves for the two governance modes. When applied to the real-life context, one may think of a firm that has to decide on a certain business function, such as R&D, either to be organized by using external resources via networking or by choosing the internationalization option. Internalizing the respective business activities always offers net benefits of H. When opting for networking, the resulting net benefits are determined by the number of commitments.

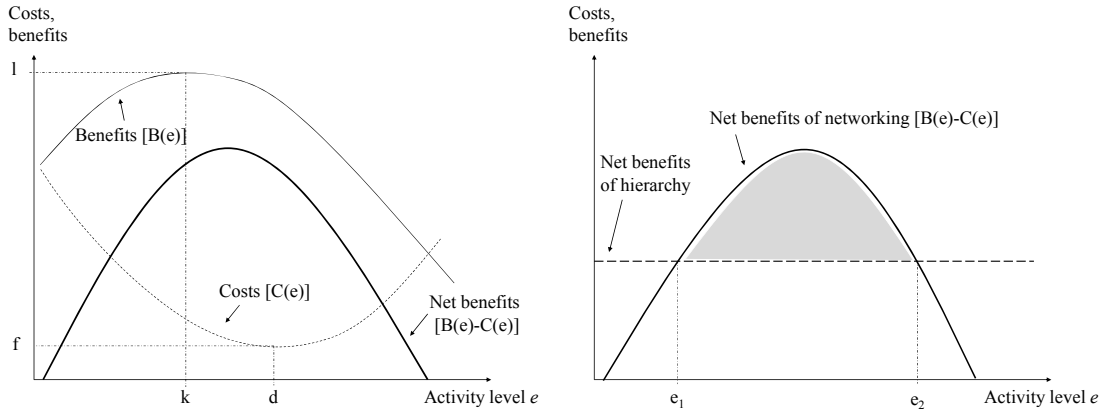

Figure 9. *Costs and Benefits of Networking*

The decision rule for organizational choice implies superiority of the network form of organization for activity levels between e_1 and e_2. Here, $B(e) - C(e) > H$ applies, which means that the benefits of networking exceed the costs of networking insofar as they cover the opportunity costs, i.e., the net benefits of hierarchy. However, for $B(e) - C(e) < H$, which corresponds to activity levels lower than e_1 and higher than e_2, interconnected firms are disadvantaged relative to those which rely on hierarchical organization. Note that we can certainly think of business functions for which the net benefits of networking always fall below the gains one may derive from the hierarchical solution. This might particularly be true for activities which concern core competencies of the firm, such as costumer acquisition, for example. Furthermore, for the sake of the argument H is defined as being constant over all activity levels. However, refining this heuristic model one could conceive other cost functions which depend on the activity level.

The simple heuristic model suggested above illustrates that economic actors might deliberately choose other organization structures than networking or might also be indifferent between alternative governance modes as a result of a cost-benefit analysis. Most importantly, non-linear developments of networking benefits and costs should crucially affect the choice of a reasonable activity level within the network form of organization.

7. Summary and Conclusion

Our theoretical analysis indicates that there is no linear additive relationship between the connectedness of a firm and the value effects of additional external resources. In conclusion, this chapter seeks to encourage a more general approach to network analysis, allowing for opportunity costs of alternatively available forms of governance and non-linearities to affect the value of additional shared and non-shared resources. It follows from this perspective that differences in rent generation may not only be driven by firm- and relation-specific factors, but particularly by networking activity. High activity levels of networking imply declining benefits due to inefficiencies in information flows and scarce management capacity which – in conjunction with restricted tie breaking capabilities – might hinder intensive exploitation of connections. Furthermore, when exceeding a certain threshold of network size, the costs of networking rise exponentially due to increasing costs resulting from coordination efforts.

Dyer, Kale and Singh (2001) report that 49% of alliances did not have established any kind of metric to measure network performance. Considering potential constraints and dysfunctionalities of networking, firms are in the danger of not recognizing rising costs from network extension which might result in a competitive disadvantage of the interconnected firm.

8. References

Ahuja, G. (2000). The Duality of Collaboration: Inducements and Opportunities in the Formation of Interfirm Linkages. *Strategic Management Journal,* 21: 317-343.

Alchian, A.A. (1950). Uncertainty, Evolution and Economic Theory. *Journal of Political Economy,* 58: 211-221.

Amit, R. and Schoemaker, P.J.H. (1993). Strategic Assets and Organizational Rent. *Strategic Management Journal,* 14: 33-46.

Arend, R.J. and Seale, D.A. (2005). Modeling Alliance Activity: An Iterated Prisoners' Dilemma with Exit Option. *Strategic Management Journal,* 26: 1057-1074.

Argote, L., McEvily, B. and Reagans, R. (2003). Managing Knowledge in Organisations: An Integrative Framework and Review of Emerging Themes. *Management Science,* 49: 371-392.

Axelrod, R. (1984). *The Evolution of Cooperation.* New York: Basic Books.

Barney, J.B. (1991). Firm Resources and Sustained Competitive Advantage. *Journal of Management*, 17: 99-120.

Baum, J.A.C. and Oliver, C. (1991). Institutional Linkages and Organizational Mortality. *Administrative Science Quarterly,* 36: 187-218.

Burt, R.S. (1992). *Structural Holes: The Social Structure of Competition***.** Cambridge, MA: HUP.

Burt, R.S. (2005). *Brokerage and Closure: An Introduction into Social Capital.* New York, NY: Oxford University Press.

Burt, R.S. and Knez, M. (1995). Kinds of Third-Party Effects on Trust. *Rationality and Society,* 7: 255-292.

Camerer, C. and Knez, M. (1997). Coordination in Organizations: A Game Theoretic Perspective. In Shapira, Z. (Ed.) *Organizational Decision Making.* Cambridge, MA: Cambridge University Press.

Cohen, W. and Levinthal, D. (1990). Absorptive Capacity: A New Perspective on Learning and Innovation. *Administrative Science Quarterly,* 35: 128-152.

Coleman, J.S. (1990). *Foundations of Social Theory.* Cambridge, MA: Harvard University Press.

Conner, K.R. (1991). A Historical Comparison of Resource-Based Theory and Five Schools of Thought within Industrial Organization Economics: Do We Have a New Theory of the Firm? *Journal of Management,* 17: 121-154.

Contractor, F.J. and Lorange, P. (1988). Why Should Firms Cooperate? The Strategy and Economics Basis for Cooperative Ventures. In Contractor, F.J. and Lorange, P. (Eds.) *Cooperative Strategies in International Business.* New York: Heath.

Davis, L.S. (2006). Growing Apart: The Division of Labor and the Breakdown of Informal Institutions. *Journal of Comparative Economics,* 34: 75-91.

Dore, R. (1983). Goodwill and the Spirit of Market Capitalism. *British Journal of Sociology,* 34: 459-482.

Doz, Y. (1988). Technology Partnerships between Larger and Smaller Firms: Some Critical Issues. In Contractor, F.J. and Lorange, P. (Eds.) *Cooperative Strategies in International Business,* New York: Heath.

Dyer, J.H. (1997). Effective Interfirm Collaboration: How Firms Minimize Transaction Costs and Maximize Transaction Value. *Strategic Management Journal,* 18: 535-556.

Dyer, J.H., Kale, P. and Singh, H. (2001). How to Make Strategic Alliances Work. *MIT Sloan Management Review,* 42: 37-43.

Dyer, J.H. and Singh, H. (1998). The Relational View: Cooperative Strategy and Sources of Interorganizational Competitive Advantage. *Academy of Management Review,* 23: 660-679.

Eisenhardt, K. and Schoonhoven, C. (1996). Resourced Based View of Strategic Alliance Formation: Strategic and Social Effects in Entrepreneurial Firms. *Organization Science,* 7: 136-150.

Frank, R.H. (2000). *Microeconomics and Behavior – Fourth Edition.* Boston: McGraw-Hill.

Gargiulo, M. and Benassi, M. (1999). The Dark Side of Social Capital. In Gabbay, S. and Leenders, R. (Eds.) *Social Capital and Liability.* Norwell, MA: Kluwer.

Granovetter, M. (1973). The Strength of Weak Ties. *American Journal of Sociology,* 78: 1360-1380.

Granovetter, M. (2005). The Impact of Social Structure on Economic Outcomes. *Journal of Economic Perspectives,* 19: 33-50.

Gulati, R. (1995). Social Structure and Alliance Formation Patterns: A Longitudinal Analysis. *Administrative Science Quarterly,* 40: 619-652.

Gulati, R. and Gargiulo, M. (1999). Where do Interorganizational Networks Come From? *American Journal of Sociology,* 104: 1439-1494.

Gulati, R., Lawrence, P.R. and Puranam, P. (2005). Adaption in Vertical Relationships: Beyond Incentive Conflict. *Strategic Management Journal,* 26: 415-440.

Gualti, R. and Singh, H. (1998). The Architecture of Cooperation: Managing Coordination Costs and Appropriation Concerns in Strategic Alliance. *Administrative Science Quarterly,* 43: 781-814.

Hamel, G. (1991). Competition for Competence and Interpartner Learning within International Strategic Alliances. *Strategic Management Journal,* Summer Special Issue, 12: 83-103.

Hart, O. (1995). *Firms, Contracts and Financial Structure.* Oxford: Clarendon Press.

Hennart, J. (1991). The Transaction Cost Theory of Joint Ventures: An Empirical Study of Japanese Subsidiaries in the United States. *Management Science,* 37: 483-497.

Henrich, J. (2004). Cultural Group Selection, Coevolutionary Processes and Large-Scale Cooperation. *Journal of Economic Behavior and Organization,* 53: 3-35.

Henrich, J. and Henrich, N. (2006). Culture, Evolution, and the Puzzle of Human Cooperation. *Cognitive Systems Research,* 7: 220-245.

Homans, G. (1961). *Social Behavior: Its Elementary Forms.* New York: Harcourt Brace.

Jarillo, J.C. (1988). On Strategic Networks. *Strategic Management Journal,* 9: 31-41.

Kali, R. (2003). Business Groups, the Financial Market and Modernization. *Economics of Transition,* 11: 671-696.

Klein, B., Crawford, R. and Alchian, A.A. (1978). Vertical Integration, Appropriable Rents, and the Competitive Contracting Process. *Journal of Law and Economics,* 21: 297-326.

Kogut, B. and Zander, U. (1992). Knowledge of the Firm, Combinative Capabilities, and the Replication of Technology. *Organization Science,* 3: 383-397.

Labianca, G. and Brass, D.J. (2006). Exploring the Social Ledger: Negative Relationships and Negative Asymmetry in Social Networks in Organizations. *Academy of Management Review,* 31: 596-614.

Lavie, D. (2006). The Competitive Advantage of Interconnected Firms: An Extension of the Resource-Based View. *Academy of Management Review,* 31: 638-658.

Lorenzen, M. (2002). Ties, Trust and Trade: Elements of a Theory of Coordination in Industrial Clusters. *International Studies of Management and Organization,* 31: 14-34.

Nahapiet, J. and Ghoshal, S. (1998). Social Capital, Intellectual Capital, and the Organizational Advantage. *Academy of Management Review,* 22: 242-266.

Oliver, C. (1997). Sustainable Competitive Advantage: Combining Institutional and Resource-Based Views. *Strategic Management Journal,* 18: 697-714.

Park, S.H. and Ungson, G.G. (2001). Interfirm Rivalry and Managerial Complexity: A Conceptual Framework of Alliance Failure. *Organization Science,* 12: 37-53.

Parkhe, A. (1991). Interfirm Diversity, Organizational Learning, and Longevity in Global Strategic Alliances. *Journal of International Business Studies,* 22: 579-601.

Parkhe, A. (1993). The Structuring of Strategic Alliances: A Game Theoretic and Transaction Cost Examination of Interfirm Cooperation. *Academy of Management Journal,* 36: 794-829.

Parkhe, A., Wasserman, S. and Ralston, D.A. (2006). New Frontiers in Network Theory Development. *Academy of Management Review,* 31: 560-568.

Pillai, K.G. (2006). Networks and Competitive Advantage: A Synthesis and Extension. *Journal of Strategic Marketing,* 14: 129-145.

Pittaway, L., Robertson, M., Munir, K., Denyer, D. and Neely A. (2004). Networking and Innovation: A Systematic Review of the Evidence. *International Journal of Management Reviews,* 5/6: 137-168.

Podolny, K.L. and Page, J.M. (1998). Network Forms of Organization. *American Review of Sociology,* 24: 57-76.

Powell, W.W. (1990). Neither Market nor Hierarchy: Network Forms of Organization. In Staw, B. and Cummings, L.L. (Eds.) *Research in Organizational Behavior Vol. 12*. London: JAI.

Powell, W.W., Koput, K.W. and Smith-Doerr, L. (1996). Interorganizational Collaboration and the Locus of Innovation: Networks of Learning in Biotechnology. *Administrative Science Quarterly,* 41: 116-145.

Ring, P.S. and Van de Ven, A.H. (1992). Structuring Cooperative Relationships Between Organizations. *Strategic Management Journal,* 13: 483-498.

Riordan, M.H. and Williamson, O.E. (1985). Asset Specificity and Economic Organization. *International Journal of Industrial Organization,* 3: 365-378.

Rowley, T.J. (1997). Moving Beyond Dyadic Ties: A Network Theory of Stakeholder Influences. *Academy of Management Review,* 22: 887-910.

Sakakibara, M. (1997). Heterogeneity of Firm Capabilities and Cooperative Research and Development: An Empirical Examination of Motives. *Strategic Management Journal,* Summer Special Issue, 18: 143-164.

Sako, M. (1992). *Prices, Quality and Trust: Inter-Firm Relations in Britain and Japan.* Cambride, MA: Cambridge University Press.

Shortell, S. and Zajac, E. (1988). Internal Corporate Joint Ventures: Development Processes and Performance Outcomes. *Strategic Management Journal,* 9: 527-542.

Uzzi, B. (1996). The Sources and Consequences of Embeddedness for the Economic Performance of Organizations: The Network Effect. *American Sociological Review,* 61: 674-698.

Uzzi, B. (1997). Social Structure and Competition in Interfirm Networks: The Paradox of Embeddedness. *Administrative Science Quarterly,* 42: 35-67.

Varian, H.R. (1997). *Intermediate Microeconomics – Fourth Edition.* New York, NY: Norton.

Wasserman, S. and Faust, K. (1994). *Social Network Analysis.* Cambridge, MA: Cambridge University Press.

White, S. and Siu-Yun Lui, S. (2005). Distinguishing Costs of Cooperation and Control in Alliances. *Strategic Management Journal,* 26: 913-932.

Williamson, O.E. (1985). *The Economic Institutions of Capitalism.* New York, NY: Free Press.

Williamson, O.E. (1991). Comparative Economic Organization: The Analysis of Discrete Structural Alternatives. *Administrative Science Quarterly,* 36: 269-296.

Williamson, O.E. (1996). *The Mechanisms of Governance.* New York: Oxford University Press.

Windsperger, J. (1987). Zur Methode des Transaktionskostenansatzes. *Zeitschrift für Betriebswirtschaft,* 57: 59-76.

Witt, P. (2004). Entrepreneurs' Networks and the Success of Start-Ups. *Entrepreneurship & Regional Development*, 16: 391-412.

Zajac, E.J. and Olsen, C.P. (1993). From Transaction Cost to Transactional Value Analysis: Implications for the Study of Interorganizational Strategies. *Journal of Management Studies,* 30: 131-145.

III. THE IMPACT OF E-WORD-OF-MOUTH ON THE SALES DISTRIBUTION IN ONLINE BOOK RETAILING*

1. Abstract

Exploring the long tail phenomenon, we empirically analyze whether online reviews, discussion forums, and product recommendations reduce search costs and actually alter the sales distribution in online book retailing towards equality. By developing an innovative approach, we provide the first long tail conversion model for the German online market, based on publicly available sales data. Analyzing a data set containing more than 30,000 different books, we find out that e-Word-of-Mouth reduces search costs by facilitating the identification and quality assessment of adequate books. Yet, unconditional quantile regressions reveal differences in the functionality of eWOM with respect to popular and unpopular books.

* This chapter originates from joint work with Thomas Ehrmann. An early version of this chapter was presented at the ICIS 2008 Conference in Paris, F. See: Ehrmann, T. and Schmale, H. (2008), The Hitchhiker's Guide to the Long Tail: The Influence of Online Reviews and Product Recommendations on Book Sales – Evidence from German Online Retailing. Proceedings of the International Conference on Information Systems, 2008, Paris, F.

2. Introduction

"All truths are easy to understand once they are discovered – the trick is to exploit them."
Guy Kawasaki

While big brick-and-mortar book retailers' assortments comprise a maximum of approximately 200,000 books, online retailers like "Amazon" generate more than a quarter of their sales volume with titles which are positioned beyond the top 200,000 in their sales ranking lists. Hence, it could be argued that the real existing book market exceeds the conventional brick-and-mortar market by one-third in sales volume. The phenomenon of a high total sales volume of products which individually sell a very small number of copies is known as "the long tail", a term coined by Chris Anderson, chief editor of the Wired-Magazine (Anderson, 2006).

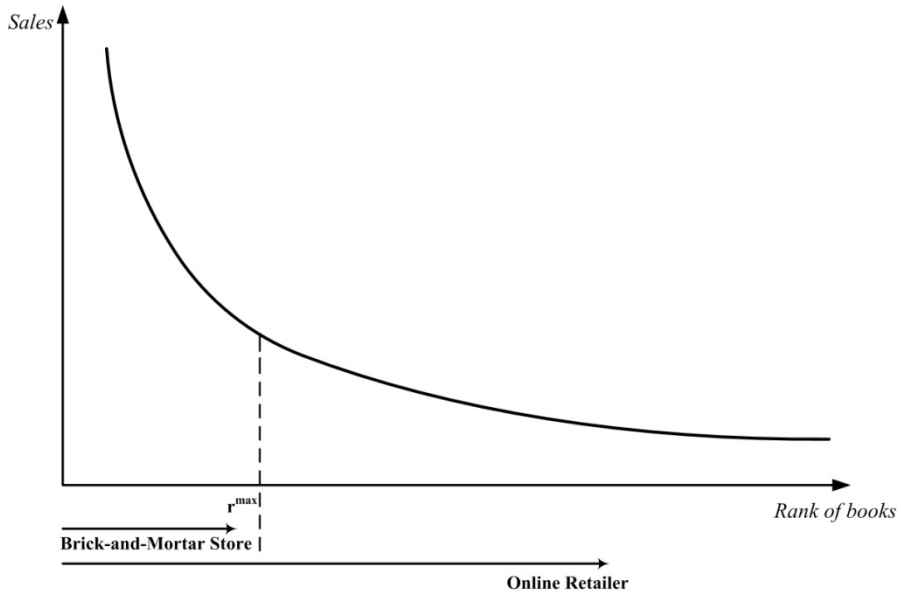

Figure 10. *The Long Tail of Books*

Traditionally, the world of retailing followed the Pareto principle (Brynjolfsson, Hu and Simester, 2006). The Pareto principle, also known as 80/20-rule, states that 80 percent of total sales are generated by 20 percent of the offered products. Consequently, markets were typically dominated by a winner-take-all mentality. However, the long tail phenomenon suggests that those products that sell too little to be accommodated in brick-and-mortar stores are responsible for a much higher fraction of total sales than predicted by the Pareto principle. Consequently, the

question is whether it is profitable for the supply side to move focus away from hits, i.e. the "short head", towards less popular and more obscure products, i.e. the long tail.

Both supply-side and demand-side drivers affect the profitability of a long tail strategy. On the one hand, online markets feature – as opposed to conventional stores – nearly unlimited shelf space (Bianco, 1997; Brynjolfsson, Hu and Smith, 2003). From this follows that the opportunity costs of listing obscure products converge to zero, eventually making it profitable to list niche products (Borenstein and Saloner, 2001). Furthermore, online retailing allows to aggregate geographically dispersed demand, which has positive effects on the overall sales volume of low-selling products (Lyster, 1999). On the other hand, the long tail phenomenon is driven by changes in consumer behavior. The implementation of powerful search tools and automated recommender systems entails declining consumer search costs in terms of opportunity costs of time. Thus, it might become more advantageous for consumers *not* to buy well-known and prominently positioned hit products, but to expand search effort in order to find products which better match personal requirements and preferences than the mainstream product (e.g. Broniarczyk, Hoyer and McAlister, 1998; Senecal and Nantel, 2004).

Long tail theory's prediction with respect to the market is that obscure products, which lack the high awareness of hit products among buyers, will profit disproportionately high from search cost reductions. Hence, declining consumer search costs will even out demand between popular and niche products, thereby leading to reduced demand inequity, and to a further flattening of the sales distribution (Brynjolfsson, Hu and Smith, 2003; Brynjolfsson et al., 2006).

In recent years, researchers have begun to reveal the determinants of long tail formation in more detail (e.g. Brynjolfsson et al., 2006; Elberse and Oberholzer-Gee, 2006; Oestreicher-Singer and Sundararajan, 2006). In particular, and most interesting to our research, scholars and practitioners emphasized the relevance of user-generated online reviews and discussion forums for driving demand down the tail by reducing product quality uncertainty (e.g. Kawasaki, 2006). However, despite this widespread belief there still is only sparse empirical evidence on the effective-

ness of e-Word-of-Mouth (eWOM) instruments for reducing search costs in the long tail, and particularly, on their role in changing the sales distribution.

In this study, we empirically analyze the effects of online reviews, discussion forums, and automated product recommendations on the demand of individual products and the overall distribution of sales. Using weekly data of over 30,000 books from Germany's biggest online book retailer, Amazon.de, we fit unconditional quantile regression models (Firpo, Fortin and Lemieux, 2009) in order to explore the effects of eWOM on the sales distribution. This new regression method allows for evaluating the impact of explanatory variables on the unconditional distribution of an outcome variable and thus permits to draw individual conclusions with respect to the influence of these variables on the long tail and the short head, respectively.

One major challenge in analyzing the determinants of long tail formation is to estimate actual demand levels from corresponding sales ranks, since proprietary sales data typically is hard to obtain. However, using an innovative approach based on publicly available sales data, we provide the first estimation of a long tail conversion model inferring actual demand levels for obscure products from corresponding sales ranks for the German book market.

The remainder of this chapter is structured as follows. First, we will give an overview of the theoretical background and introduce our research questions as well as our hypotheses. Second, we describe the data. Third, we estimate the long-tail conversion model. Fourth, we introduce the methodology as well as the model specifications and present the results of our empirical tests. Finally, we discuss our key results and conclude.

3. Theoretical Background and Hypotheses

3.1 Search Costs and the Long Tail

Analyzing the U.S. VHS and DVD market Elberse and Oberholzer-Gee (2006) show that a supply-sided broadening of the product range indeed comes along with an increasing popularity of niche products. Nevertheless, such broadening also implies a very high number of titles which rarely sell at all. This might constitute a non-negligible expense factor for retailers pursuing a long tail strategy. Hence, another fundamental prerequisite for the successful implementation of a long tail strategy – other than broadening the assortment to a maximum diversified product range – is the reduction of demand-side search costs (Anderson, 2006; Chernev, 2006; van Herpen and Pieters, 2007).

First empirical evidence indicates that search costs decline in the internet channel compared to conventional channels like catalogue retailing (Brynjolfsson et al., 2006). Interpreting product search as a two-staged process of (1) identification of products and (2) assessment of fit between the product's characteristics and a consumer's quality and functional requirements (Stiglitz, 1989), there are two main classes of setting levers for reducing search costs.

First, search costs emanating from attempts to identify potential products can be reduced by developing specialized search filters and automated recommender systems which provide personalized information to customers (Ansari, Essegaier and Kohli, 2000; Häubl and Trifts, 2000). For instance, Amazon provides on most of the books' web pages information on other books that have been co-purchased by other customers. Following these product recommendation links, customers possibly locate formerly unknown but adequate products with a higher probability than by random search, thereby incurring lower search costs due to less total time spent searching. Oestreicher-Singer and Sundararajan (2006) provide empirical evidence on the influence of Amazon's co-purchase network on product sales concentration. Relatedly, Chen, Wu and Yoon (2004) find that product recommendations improve sales more for less popular books than for more popular books at Amazon.com.

Second, search costs can be reduced by implementing instruments which are targeted toward reducing quality uncertainty. As described above, product recommendations foster demand for niche products. However, a prominent position in the co-purchase network necessitates that a high number of customers previously bought the product in a bundle with a more prominent good. Actually, most of the niche products are per definition not rated by the purchase behavior of the broad customer mass. As Guy Kawasaki (2006: 1) puts it: "this is where two cool concepts butt head: long tail vs. wisdom-of-the-crowds." On the contrary, online customer reviews can support the assessment of obscure products by providing more detailed information on the products' characteristics even though very few copies have previously been sold. That is to say, one sole antecedent purchase and review post can potentially help to reduce the perceived risk of wasting time and money on inadequate products (Anderson, 2006; Bakos, 1998; Kawasaki, 2006; Oestreicher-Singer and Sundararajan, 2006). With regard to a long tail strategy, reviews are of high economic relevance due to their cost efficiency, since this form of community content requires relatively low maintenance compared to retailer self-provided product information.

Among long tail scholars, there is widespread belief that with respect to the demand side of the market, decreasing transaction costs come along with an increasing heterogeneity in consumption patterns (e.g. Brynjolfsson et al., 2006; Hervas-Drane, 2007). Theory argues that consumers benefit from increasing product variety in being able to buy products that better match personal preferences than conventional mainstream products. As the number of easily accessible offerings increases, it becomes more likely that adequate products will be found (e.g. Broniarczyk et al., 1998; Chernev, 2003).[29] If this is true, the Pareto sales distributions of brick-and-mortar markets reflect – the long tail view holds – suboptimal choices from an individual perspective to some extent. Hence, it is the "consumers' appetite for niche products" (Elberse and Oberholzer-Gee, 2006: 4) which constitutes a prerequisite for a lucrative long tail to emerge.

Based on long tail theory's basic assumptions, Brynjolfsson et al. (2006) develop a theoretical model that illustrates the effects of a demand-side search cost reduction on sales concentration.

[29] Moreover, theory argues that consumers benefit from large assortments through increased choice flexibility in the face of changing preferences and improved support of variety-seeking behavior (Boyd and Bahn, 2009; Chernev, 2006; Kahn and Lehmann, 1991).

Key to the model is the fact that products differ *ex ante* in search costs. On the one hand, "key" products which are heavily promoted by expensive marketing campaigns generate lower search costs since these products entail a high awareness among consumers. On the other hand, purchasing niche products involves high search costs because these products are not visible per se. The assumption of heterogeneous search costs implies a change in sales dispersion as a result of a search cost reduction. In contrast, if products were assumed to be homogeneous, a reduction in search costs should affect all products equally, implying a stable sales concentration.

In our research setting, we allow for the fact that products might differ ex ante in search costs. Consider for example products that stand out of the broad mass due to the fact that they have been reviewed in TV shows or have been dragged into the spotlight by literary awards. In contrast, niche products are commonly not subject to heavy advertising or promotion campaigns, and consequently, there are only few, if any, other sources of information available. Consequently, user generated online reviews – or e-Word-of-Mouth generally – should be especially influential on the sales of these obscure products since this is the only source of information available (Chevalier and Mayzlin, 2006).

3.2 Research Questions and Hypotheses

We analyze whether user generated online reviews, discussions and recommendations can help to reduce search cost and whether these instruments in fact foster a shift in consumption patterns away from hits to the niche. If this is the case, we should observe non-uniform effects – in the sense of strength or direction of influence – of those instruments on sales across books with respect to their ranking. In particular, we pose following global research questions:

(1) Do customer reviews, discussion forums and product recommendations affect sales in general?

(2) Do customer reviews, discussion forums and product recommendations alter the sales distribution towards equality?

Some empirical evidence has been provided by scholars on the basic influence of online customer reviews, or generally, eWOM, on sales (e.g. Chevalier and Mayzlin, 2006). On the one hand, eWOM can create customer awareness and on the other hand it may be one of the only sources of information about the quality of experience goods (Li and Hitt, 2004; Senecal and Nantel, 2004), i.e. products which cannot be assessed prior to consumption (Akerlof, 1970). Hence, as customers face high risk to waste money and time on inadequate products, they might rely on other customers' prior experiences in order to reduce quality uncertainty.

However, customer ratings as a measure of product quality have been controversially discussed in the literature. For example, Chevalier and Goolsbee (2003a; 2003b), Chevalier and Mayzlin (2006), Li and Hitt (2004) and Dellarocas, Awad and Zhang (2004) show that the average rating – Amazon's review system allows for customer ratings on a five-point scale – significantly affects product sales. In contrast, Chen et al. (2004) and Liu (2006) prove positive influence by the number of reviews posted rather than the reviews' valences. Hu, Pavou and Zhang (2007) ascribe this inconsistency of findings with regard to review valence to the fact that the average rating might be a skewed measure for quality. They show that reviews are not normally distributed over the five-point scale but j-shaped, implying a very high number of positive reviews, a smaller number of negative reviews, and an even smaller number of medium reviews. One possible explanation for the low number of medium ratings is the under-reporting bias, which implies that review posting requires a certain involvement by the customer. This involvement in turn is higher when a product is either really appreciated or disliked. The high number of positive reviews can be attributed to a purchasing bias. Most reviewers ex-ante seem to have a positive attitude towards the product they review, because they have obviously purchased it. As a consequence, customers might display a negativity bias in assigning higher credibility to negative rather than positive reviews because of their rare occurrence (Sen and Lermann, 2007). With regard to the literature, we try to disentangle the effects of review valence by explicitly accounting for differences in the influence of negative and positive reviews. Hence, with respect to review valence we pose following hypotheses:

Hypothesis 1a. A high fraction of good reviews positively influences the title's sales.

Hypothesis 1b. A high fraction of bad reviews negatively influences the title's sales.

Because of its experience good character buying a book can constitute a rather complex situation. As noted above, people might index their expectations regarding fit and quality of the book to consumer reviews. However, the interpretation of feedback and the assessment of its usefulness in many cases require a certain amount of prior knowledge on the topic or the author, for instance. Moreover, as the consumption of books is also characterized by an affective and sensory experience of aesthetic or sensual pleasure (Hirschman and Holbrook, 1982) – i.e. reading targets on satisfying emotional wants – reviews are subjective and thus depend on the individual emotional constitution of the reviewer. As this constitution can hardly be assessed by potential buyers, it is difficult to appraise whether the evaluation applies to oneself (Sen and Lermann, 2007). Consequently, people might try to cope with this situation by transforming the complex context into a rather simple decision rule. In particular, they might orientate themselves to the number of review posts as an indicator for reviewer involvement and popularity of the book. However, it is natural to assume that the magnitude of the positive impact of additional reviews might be decreasing for higher numbers of reviews.

Hypothesis 2a. A high number of customer reviews positively influence the title's sales.

Hypothesis 2b. The magnitude of the positive impact is decreasing for higher numbers of reviews.

Another possibility to cope with uncertainty in terms of feedback credibility and applicability is to rely on opinion leaders. Since top reviewers possess an above-average history of helpful reviews, customers might impute certain knowledge to them which comes along with higher credibility. Moreover, books often have a symbolic character, i.e. people read books in order to appeal cultured or literate. Hence, people face the risk of reading the "wrong" books and consequently might prefer books which top reviewers have found worth reading:

Hypothesis 3. Being reviewed by a top reviewer positively influences the title's sales.

Due to the social nature of people books play an important role as means of social interaction. That is, people seem to be devoted to hear the same music, watch the same movies and read the same magazines and books (Elberse and Oberholzer-Gee, 2006). One theoretical explanation for this fact is that utility derived from e.g. reading – at least to some extent – depends on the ability to appreciate it. This ability in turn improves with the accumulation of author or topic specific *consumption capital*. Put differently: The more you know, the more you appreciate it (Stigler and Becker, 1977). Consumption capital thereby can be accumulated either by consuming books of the same author or of related topics or by discussions with others. Since it is costly to search for someone to interact with, people might prefer reading books with visible 'user communities' because it becomes more likely to find interaction partners. Furthermore, large user communities indicate high involvement and thus might convince that the book is worth reading. Hence, with respect to the relationship between discussions and sales we pose the following research hypothesis:

Hypothesis 4. Being discussed positively influences the title's sales.

Providing community content is one strategy to reduce customer uncertainty involved in buying experience goods. Another more and more frequently pursued strategy is to enable free trial. Specifically, retailers provide reading excerpts (e.g. the first twenty pages of a book). The excerpts allow for relatively unbiased quality assessment by the customer und thus might contribute to search cost reduction:

Hypothesis 5. Featuring a reading excerpt positively influences a title's sales.

As already noted, automated recommender systems can reduce search costs by simplifying the identification of adequate books. In this respect, a book's network position on a retailer's e-commerce site influences the amount of traffic which is being directed to the product's detail page. The network position thereby is determined by the books it links to, and those that link to it. However, books that are either linked to by one or more popular books or by a high number of moderate-sellers are likely to enjoy an increase in sales on account of improved customer

traffic. Put differently, the book should "enter" more potential buyers' individual choice sets. Besides that, product links do not only entail higher traffic levels but they might also provide implicit content information because they allow customers to relate unknown books to something they possibly know. Hence with respect to automated recommendation links we deduce following hypothesis:

Hypothesis 6. Being exposed to high network influence positively impacts the title's sales.

As discussed in the previous section, long tail theory argues that reducing search costs fosters a shift in sales concentration – away from hits to the niche. The rationale behind this is the fact that books are not homogeneous with respect to search cost. In particular, hits feature a higher awareness among customers due to the general availability of various information sources such as bestseller lists, newspaper reviews, advertising and so on. Consequently, community content (i.e. reviews and discussions), recommendations and excerpts should play a much more important role for sales success in the long tail where search costs are high. Hence we deduce:

Hypothesis 7. Community content (i.e. reviews and discussions), recommendations and reading excerpts have greater impact on sales in the tail of the sales distribution.

And, relatedly:

Hypothesis 8. Community content (i.e. reviews and discussions), recommendations and reading excerpts alter the sales distribution towards equality.

4. Data and Summary Statistics

Using Perl-based scripts, we collected individual characteristics, review data, and network information of initially 56,744 books at the public website of Amazon.de, Germany. Since we did not have access to Amazon's real sales data, we tried to select a representative sample as follows:[30]

First, we chose Amazon's bestseller lists – each main category has its own hit list – as a starting point for our data collection in order to make sure that the sample contains a sufficient amount of popular titles. Hence, we collected 9,600 titles across all 22 main categories. In order to dig deeper into the long tail, we furthermore added a random generator based sample of 10,000 books from the German "Books in Print" directory.[31] Since a large fraction of sales are concentrated in a small fraction of books (Greco, 1997), the probability of extracting niche titles was very high. However, we also wanted to collect information on product recommendations, but the co-purchase link network is a directed graph, i.e., it is possible to collect all the subsequent books one title refers to but it is not possible to identify all precedent books which refer to a title. Thus, we took the following approach to draw our main sample: Amazon.de provides up to six co-purchase links on every book's detail page (see for example book A and B in Figure 11).

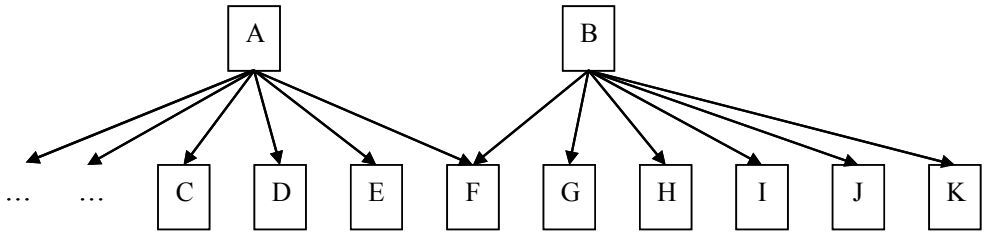

Figure 11. *Examples of Co-Purchases of Books A and B*

We added all co-purchase links of our initial sample of 19,600 books thereby broadening the sample size to 56,744 different titles with co-purchase information (books C to K in Figure 11). In addition to this we collected following data for each of the books in our sample: isbn (unique

[30] This systematic conforms to the common method of data collection (e.g. Chevalier and Mayzlin, 2006).
[31] See "Verzeichnis lieferbarer Buecher", www.vlb.de.

serial number), title, author, price, ship time, release date, format, publisher, and a set of different review variables.

Furthermore, we tried to control for other sources of "visibility" which could explain the variation in sales: (1) a dummy indicating whether a book appeared in one of the three leading German literature TV shows in the month before data collection, (2) a dummy indicating whether an author received a literary prize within the last three years ("Georg Büchner Prize", "Ingeborg Bachmann Prize", "Peace Prize of the German Book Trade", "Noble Prize in Literature", "Pulitzer Prize"), and (3) whether a book was published by one of the 50 biggest publishing houses as an indicator for marketing budgets.

Finally, we collected the sales rank of each book. The top-selling book at Amazon.de has a sales rank of one, the lower sellers are assigned higher sequential ranks. Inspections of data as well as buying experiments, which we have conducted, indicate that sales ranks are updated several times a day.[32] For the data collection this poses issues. Frequently changing ranks imply significant variations in sales ranks within one day. For example, a book which is ranked around 150,000 can move up to rank 20,000 as a result of a sale of one. Thus, collecting sales ranks resembles shooting at moving targets. Consequently, we collected sales rank data on each of our books twice a week in a four week period between March 10 and April 4, 2008. We computed an average rank for each title based on eight observations which should represent a sufficient approximation of weekly sales ranks (Rosenthal, 2008). We excluded all books which were not represented by eight observations in our database. Due to missing data, our analysis is based on 268,872 observations for 33,609 different books at Amazon.de.

[32] This accounts for both high-selling books and those whose ranks are far beyond 100,000. Ranks are changing at least every two hours.

	Mean	S.D.	Min.	Max.
0. Sales ranks	108,270.500	160,938.300	1	2,710,877
1. Fraction of 5-star ratings	0.446	0.409	0	1
2. Fraction of 1-star ratings	0.039	0.125	0	1
3. Number of reviews	6.895	31.689	0	2,569
4. Top reviews (dummy)	0.195	0.396	0	1
5. Discussions (dummy)	0.013	0.111	0	1
6. Reading excerpt (dummy)	0.158	0.365	0	1
7. Link value	19.279	88.109	0	3,884.360
8. Price	15.787	14.342	2	980
9. Ship time	0.422	1.490	0	35
10. Paperback (dummy)	0.663	0.473	0	1
11. Weeks since release	188.452	203.926	0	5,639.286
12. TV appearance (dummy)	0.000	0.012	0	1
13. Top50 publisher (dummy)	0.285	0.451	0	1
14. Literary prize (dummy)	0.001	0.023	0	1

N = 33,609.

Table 4. *Descriptive Statistics*

Table 4 displays the summary statistics of our sample. Sales ranks vary between 1 and 2,710,877. 19.5 percent of titles had a rating by a top reviewer. The mean of average star rating is 4.3 (s.d. 0.79) indicating that reviews are very positive on average, which is in line with previous findings (e.g. Chevalier and Mayzlin, 2006; Hu et al., 2007). 72.4 percent of the books in our sample had at least one review posted. On average, each book has 6.9 reviews.

5. Long Tail Conversion Model for Estimating Sales

Brynjolfsson et al. (2003) and Chevalier and Goolsbee (2003a) find that the relationship between sales ranks and sales at Amazon.com is approximately log-linear constituting a power-law distribution: $\ln[sales_i] = \beta_0 + \beta_1 \ln[rank_i]$. Using proprietary sales data from a publisher, Brynjolfsson et al. (2003) calibrate the relationship between sales and sales ranks by estimating β_0=10.526 and β_1=-0.871 for Amazon.com. Most empirical work using publicly available data resorts to these parameters for estimating sales. This seems to be appropriate for studies based on the American market; however, these parameters are less suitable when analyzing consumer behavior in other countries. Here, this model would obviously overestimate sales of books, taking into account that the American market exceeds the German approximately by a factor of four. Consequently, we estimated our own conversion model, which should be more suitable for the data from Amazon.de.[33] Since we did not have access to proprietary data we took another approach: Amazon.de provides information on remainder of stock for books which they have five or less copies left. This information is updated several times a day. We extracted all the books within our database which had this information displayed and checked stock variations and corresponding sales rank variations five times a day for a seven-day period in order to compute weekly sales and sales ranks.

Our initial sample contained 14,089 books of which about 6,000 exposed variations in the remainder of stock. We excluded all titles which had positive variations in stock and all those which had less than one copy left at the end of day seven. We did this to ensure that the variation in sales ranks can be related to variations in stock, since Amazon.de does not stop selling a product which is not in stock. Further minimizing potential effects of unobserved replenishment we excluded all titles from the sample that displayed declining sales ranks but either no variation in stock or positive variation (i.e. sale and replenishment within the same 2½ hours) during

[33] In a narrow sense, the log-linear relationship between sales ranks and sales is an assumption. However, Brynjolfsson et al. (2003), Schnapp and Alwine (2001), Chevalier and Goolsbee (2003a) and Rosenthal (2008) all find that this relationship is approximately log-linear for the American book market. Hence, we follow the literature and assume that this relationship applies to the German book market as well.

the collection period. Finally, we excluded all titles with increasing ranks (i.e. no sale) and positive variation in stock (i.e. replenishment) within the collection period.

Our final sample contained 540 books with weekly sales between one and four books and average rankings between 53,193 and 959,888. The mean number of sales was 1.44 and the mean average rank was 163,322. Table 5 shows our regression results.

Method: OLS	**coefficients**	**s.e.**
Ln(average rank)	-0.656***	(0.043)
Constant	8.114***	(0.516)
N	540	
F	235.518***	
Adjusted R^2	0.30	

Dependent variable: Ln(sales). Standard errors in parentheses. Significance levels (2-tailed): ***p<0.001; **p < 0.01; *p < 0.05; †p < 0.1.

Table 5. *Long Tail Conversion Function*

We estimated β_0=8.114 and β_1=-0.656. The model is highly significant, although the R2 is not as high as expected. Rosenthal (2008) reports that Amazon.com's new ranking system includes own sales as well as third party sales, which could be an explanation for this value. A limitation of this approach is the fact that we relied on long tail data. In particular, small errors with respect to the log-normal parameters which were estimated in the tail of the curve might be magnified significantly near the head of the curve, possibly leading to wrong estimates for the top sellers. Yet, anecdotal evidence suggests that our estimates fit quite well: Welt-Online (2008) reports that the book "Wetlands" (i.e. the No.1-seller in our sample) should have sold 3,100 units per week at Amazon.de within our collection period. Our model estimates weekly sales of 3,341; this implies an accuracy of 93 percent for the top selling book. However, without having proprietary data this should be an adequate estimation of Amazon.de's sales.[34]

[34] The summary statistics for the variable sales are: mean=5.844; s.d.=29.977; min.=0.2; max.=3,440.91.

6. The Impact of Reviews and Recommendations on the Sales Distribution: Empirical Test

6.1 Methods and Model Specifications

OLS-Regression. To estimate the influence of online reviews, discussion forums, and product recommendations on sales we specified the following general form model:

(1) $$y = x'\beta+\varepsilon$$

The response variable is the actual demand of books, log sales, according to our previously introduced conversion model. The vector x of covariates is specified as follows:

First, we included a set of review variables: fraction of 5-star reviews, fraction of 1-star reviews, number of reviews, a dummy variable indicating whether a top-reviewer occurs on the books main page, and a dummy for discussion posts. Furthermore, we included a dummy for reading excerpts.

Second, we included the network's immediate influence on a book due to recommendations. This variable, link value, is a score which depends on both the number of recommendations (i.e. links) a book receives and the quality of these recommendations in terms of demand of the preceding books (Oestreicher-Singer and Sundararajan, 2006).

Finally, we considered a set of control variables including book characteristics: price, ship time, format (paperback dummy), number of weeks since the book was released, TV appearance, a Top50 publisher and a literary prize dummy.

Unconditional Quantile Regression. An important shortcoming of ordinary least squares regression – at least in this context – is that it does not allow for drawing conclusions with respect to possible changes of the sales distribution as a result of asymmetric impact of eWOM instruments. To the contrary, OLS estimates a conditional mean function that describes how the mean of Y changes with a covariate X, and thus intends to deliver the "true value" around which Y fluctuates due to an erratic component. A basic assumption here is that the error has the same

distribution whatever values may be taken by *X*. This is referred to as the pure location shift model assuming that *X* affects only the location of a conditional distribution, not the scale or other aspects of its distributional shape (Koenker and Hallock, 2001).

Yet, we actually expect eWOM to vary in impact on books across the sales distribution and, consequently, resulting estimates of OLS regressions are not necessarily indicative for the strength or direction of effects on the lower tail of the distribution.

Previous literature has commonly used conditional quantile regression models in order to measure the impact of explanatory variables on distributional statistics that go beyond the mean. Within theses models, a specified conditional quantile of the response variable is expressed as a linear function of observed covariates. Consequently, this method intends to shed light on how the entire distribution changes with certain independent variables (e.g. Koenker and Hallock, 2001).

However, despite the aforementioned shortcomings, a useful feature of OLS regressions is that they provide consistent estimates of the impact of an explanatory variable, *X*, on the population *unconditional* mean. Hence, estimates can be used to measure the impact on the mean of *Y* of e.g. increasing every observation's *X* by one unit, holding everything else constant. On the contrary, estimates obtained from conditional quantile regressions cannot be used to estimate the impact of *X* on the *unconditional* distribution. This stems from the fact that conditional quantiles do not average up to their unconditional population counterparts, which is true for conditional means.[35] Referring to the fact that conditional quantile regressions only measure the effects of the covariates on within-group dispersion (where the groups consist of observations that share the same values of the covariates *X*) Firpo, Fortin and Lemieux (2009) propose a new regression method called *unconditional quantile regression.* In addition to capturing within-group effects measured by conditional quantile regressions it also captures between-group effects.[36] This new

[35] For further details see Firpo, Fortin and Lemieux (2009).

[36] Firpo, Fortin and Lemieux (2007) give an excellent numerical example of the difference between within- and between-group effects. Consider following simple example applied to our context: Suppose reviews (e.g. measured as a binary variable) have a positive impact on log sales and the effect estimated using conditional quantile regression is stronger in the 10th than in the 90th percentile. This finding

method basically consists of running simple OLS regressions of a transformation – the recentered influence function (RIF) – of the outcome variable on the explanatory variables. The *RIF* for the quantile q_τ is formally defined as

$$RIF(Y; q_\tau) = q_\tau + \frac{\tau - I(Y \leq q_\tau)}{f_Y(q_\tau)} \quad (2)$$

with $q_\tau = Q_\tau(Y)$ being the τth population quantile of the unconditional distribution of Y. f_Y is the marginal density function of Y and I is an indicator function. The *RIF* is simply just a dichotomous variable that takes on the value $q_\tau - (1 - \tau)/f_Y(q_\tau)$ when y is below the quantile q_τ, and the value $q_\tau + \tau/f_Y(q_\tau)$ when y is above the quantile q_τ. The sample analog of the *RIF* is defined as

$$RIF(Y; \widehat{q_\tau}) = \widehat{q_\tau} + \frac{\tau - I(Y \leq \widehat{q_\tau})}{\widehat{f_Y}(q_\tau)} \quad (3)$$

where $\widehat{q_\tau}$ is the τth sample quantile and $\widehat{f_Y}$ is the kernel density estimator. Note that the mean of the *RIF* is the quantile q_τ itself.

If reviews and recommendations in fact foster a change of the sales distribution towards equality – as the long tail view suggests – we should observe the respective coefficients increase for some quantiles τ in the lower tail of the unconditional sales distribution. That is, reviews, recommendations and discussion forums should have higher impact on less popular titles (e.g. Chevalier and Mayzlin, 2006).

6.2 Empirical Results: OLS

Table 6 displays the results of the OLS regression.[37] Overall, the results indicate that the fraction of 5-star reviews, the number of reviews, discussions and reading excerpts have significantly positive effects on demand, thereby corroborating hypotheses H1a, H2a, H3, H4 and H5. This indicates that the implementation of e-Word-of-Mouth instruments is of relevance for demand.

means that reviews reduce within-group dispersion, i.e. the sales dispersion of books with reviews. However, this does not necessarily mean that increasing the rate of reviewed titles reduces overall sales dispersion with respect to the unconditional distribution. This is linked to the fact that reviews increase the conditional mean of sales of reviewed titles, which creates a gap between otherwise comparable reviewed and non-reviewed books. This is an inequality enhancing between-group effect.

[37] The examination of variance inflation factors and condition indices indicate that the assumption of non-multicollinearity is not violated (Hair et al., 1998).

The quadratic number of reviews term is significant and negative, as expected, but of negligible impact.[38] The results also support H6 as books which receive more and qualitatively better recommendations by other books in the network have higher sales. As expected, H1b is supported by the data as the fraction of 1-star reviews significantly suppresses sales.

Method: OLS	**coefficients**	**s.e.**
Fraction of 5-star ratings	0.416***	(0.009)
Fraction of 1-star ratings	-0.226***	(0.026)
Number of reviews	0.010***	(0.001)
Number of reviews^2	-0.000***	(0.000)
Top reviews (dummy)	0.130***	(0.011)
Discussions (dummy)	0.490***	(0.051)
Reading excerpt (dummy)	0.088***	(0.011)
Ln(link value)	0.376***	(0.004)
Price	-0.006***	(0.001)
Ship time	-0.044***	(0.003)
Paperback (dummy)	-0.052***	(0.010)
Weeks since release	-0.000***	(0.000)
TV appearance (dummy)	1.873***	(0.251)
Top50 publisher (dummy)	0.076***	(0.008)
Literary prize (dummy)	0.426**	(0.170)
Constant	0.307***	(0.020)
N	33,609	
F	1634.22***	
R^2	0.54	

Dependent variable: ln[sales]. Significance levels (2-tailed): ***p<0.001; **p < 0.01; *p < 0.05; †p < 0.1. Robust standard errors in parenthesis.

Table 6. *Empirical Results OLS Regression*

[38] Hair et al. (1998) propose different methods for measuring curvilinear effects, i.e. transformations of the variables and the inclusion of polynomials. We chose a quadratic term in order to allow for a straight forward interpretation of the results with respect to the effects of adding additional reviews.

Turning to the control variables: Table 6 shows that the variables price, ship time, paperback format and age of a title have significant negative impacts on sales. A book's TV appearance has a positive effect on sales as well as the fact whether a title was published by a big publisher. Books from award winning authors also generate higher sales, as expected.

6.3 Empirical Results: Unconditional Quantile Regression

Table 7 shows the results of our unconditional quantile regressions for quantiles ranging from τ=0.10 to τ=0.90.

With respect to the influence of customer reviews on the sales distribution, we obtain mixed results. In particular, positive reviews are highly influential in the tail of the sales distribution, and the positive impact of 5-star reviews decreases from the lower tail to the middle of the sales distribution. However, this development is reversed from the 40^{th} quantile on, i.e. the positive impact of 5-star reviews increases. Yet, from the 80^{th} to 90^{th} quantile the coefficient drops to its minimum (coeff.=0.357, s.e.=0.028).

Conversely, negative reviews loose in impact with ascending sales ranks (i.e. in the long tail), as the coefficient of 1-star reviews indicates. Moreover, in the lowest quantile, the coefficient changes sign indicating that even a bad review is better than having none in terms of sales. In the mid and top quantiles, negative reviews significantly suppress sales, as expected.

However, as reported in the summary section, reviews are very positive on average (mean of star rating: 4.3). Hence, the findings with regard to the effects of positive reviews might indicate that increasing the proportion of reviewed books fosters a shift towards more equal sales in the long tail but an increase in the sales concentration in the mid quantiles. In order to measure the effect of increasing the proportion of reviewed books, holding everything else constant, we ran further regressions omitting the review variables and therefore including a review dummy. The results basically corroborate the u-shaped development.

Method: RIF-Regression	**Q10**	**Q20**	**Q30**	**Q40**	**Q50**	**Q60**	**Q70**	**Q80**	**Q90**
Fraction of 5-star ratings	0.406***	0.375***	0.373***	0.397***	0.421***	0.466***	0.509***	0.514***	0.357***
	(0.013)	(0.010)	(0.010)	(0.010)	(0.011)	(0.013)	(0.017)	(0.022)	(0.028)
Fraction of 1-star ratings	0.155**	0.020†	-0.043	-0.109**	-0.157***	-0.222***	-0.320***	-0.535***	-0.677***
	(0.048)	(0.037)	(0.034)	(0.033)	(0.034)	(0.038)	(0.046)	(0.060)	(0.076)
Number of reviews	0.000***	0.001***	0.002***	0.002***	0.003***	0.006***	0.009***	0.016***	0.027***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.002)
Number of reviews^2	-0.000*	-0.000**	-0.000**	-0.000**	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Top reviews	0.085***	0.100***	0.113***	0.125***	0.142***	0.174***	0.182***	0.233***	0.184***
	(0.009)	(0.009)	(0.009)	(0.010)	(0.012)	(0.015)	(0.019)	(0.028)	(0.040)
Discussions	-0.037	0.008	0.056*	0.104**	0.132***	0.224***	0.373***	0.724***	1.420***
	(0.024)	(0.022)	(0.025)	(0.030)	(0.037)	(0.050)	(0.072)	(0.113)	(0.189)
Reading excerpt	0.112***	0.087***	0.082***	0.082***	0.077***	0.084***	0.069***	0.078**	0.053
	(0.014)	(0.011)	(0.011)	(0.012)	(0.013)	(0.015)	(0.019)	(0.026)	(0.035)
Ln(link value)	0.159***	0.167***	0.195***	0.235***	0.287***	0.371***	0.491***	0.652***	0.708***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.005)	(0.008)	(0.014)
Price	-0.009***	-0.006***	-0.005***	-0.005***	-0.005***	-0.005***	-0.005***	-0.005***	-0.004***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Ship time	-0.101***	-0.057***	-0.040***	-0.031***	-0.027***	-0.022***	-0.017***	-0.016**	0.004
	(0.006)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.005)	(0.006)
Paperback	-0.040**	-0.021*	-0.017†	-0.005	-0.012	-0.024†	-0.027†	-0.061**	-0.127***
	(0.013)	(0.010)	(0.010)	(0.010)	(0.011)	(0.012)	(0.015)	(0.021)	(0.028)
Weeks since release	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.001***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
TV appearance	-0.513***	-0.476***	-0.501***	-0.517***	-0.508***	-0.474**	-0.176	0.793*	4.222***
	(0.062)	(0.061)	(0.073)	(0.090)	(0.115)	(0.156)	(0.218)	(0.315)	(0.422)
Top50 publisher	0.020†	0.042***	0.066***	0.072***	0.081***	0.095***	0.108***	0.138***	0.128***
	(0.011)	(0.008)	(0.008)	(0.009)	(0.010)	(0.012)	(0.015)	(0.021)	(0.029)
Literary prize	0.039	0.114	0.196*	0.121	0.291*	0.308†	0.606*	1.637***	0.681
	(0.082)	(0.085)	(0.098)	(0.119)	(0.137)	(0.179)	(0.266)	(0.425)	(0.869)
Constant	-0.201***	-0.040†	0.070***	0.146***	0.225***	0.268***	0.312***	0.417***	1.043***
	(0.028)	(0.021)	(0.020)	(0.019)	(0.020)	(0.022)	(0.025)	(0.031)	(0.038)
Pseudo R^2	0.16	0.20	0.24	0.28	0.31	0.33	0.34	0.32	0.26

Dependent variable: Ln(sales). Significance levels (2-tailed): ***p<0.001; **p < 0.01; *p < 0.05; †p < 0.1. Robust standard errors in parentheses.

Table 7. *Empirical Results Unconditional Quantile Regressions*

The review dummy coefficient decreases in impact from the lower to the mid quantiles (τ=0.10: coeff.=0.478, s.e.=0.015; τ=0.40: coeff.=0.375, s.e.=0.010) and then increases from the 40th to the 70th quantile (coeff.=0.408, s.e.=0.014). Consequently, our results support theory in so far as they imply that an increase of the proportion of reviewed books fosters a reduction of sales dispersion in the long tail, however, in the mid quantiles, this has a steepening effect on the sales distribution.

Turning back to Table 7, review volume has a positive impact on sales throughout the whole distribution of book sales. Yet, in contrast to H7, which postulates lower impact of reviews in the front of the sales distribution, we see steadily increasing coefficients from quantiles τ=0.10 to τ=0.90. Hence, popularity is a more influential buying argument in the hit segment. Moreover, these results indicate that review volume does not foster a decrease but an increase of dispersion of the sales distribution.[39] The squared term is significantly negative but of negligible impact for all quantiles.

Our results do support H3 which states that reviews by opinion leaders foster sales. Top reviewers' reviews have significantly positive effects throughout the sales distribution. The impact is increasing from the lower to the high quantiles, yet in the top quantile the coefficient drops slightly.

Featuring discussion posts does not have any impact on sales in the very tail of the sales distribution, for $\tau \geq 30$ discussions have positive impact. As for review volume, this finding again contrasts H7. Furthermore, the positive impact increases monotonically from 0.056 in the 30th to 1.420 in the 90th quantile. The results indicate that discussions foster sales concentration among the more successful titles.

[39] The inference of review volume affecting sales must be tempered, though, by the possibility of endogeneity. Recent studies point to the fact that the number of reviews is not totally exogenous given that it is not only the driver of consumer purchases but can also be the outcome of sales (e.g. Duan et al., 2005). In our research setting, the relevance of this endogeneity hinges on the fact whether current sales ranks are affected by historical sales, since we use lagged review variables. Rosenthal (2008) characterizes Amazon's ranking system as being almost entirely based on the notion "what have you done for me lately" discovering that historical sales have negligible impact on decay rates, i.e. the rate a book drops in the ranking as a result of a day without a sale. Consequently, the problem of endogeneity – at least in this setting here – should be less pronounced.

Conversely, reading excerpts have a flattening effect on the sales distribution, at least in the long tail. The coefficient is positive and decreasing up to $\tau < 60$. It becomes insignificant in the top quantile.

Product recommendations have positive influence on sales overall, gaining stronger impact from the lower to the high quantiles. However, since not much is known about the exact recommendation policy of Amazon, this result has to be interpreted with caution.

Turning to the controls, price is significant and of negative sign as expected. Ship time is significant and of negative sign for $\tau < 90$. Age of the book is negative and significant but of negligible impact. Format is of negative sign, but insignificant in the mid quantiles. Interestingly, a books TV-appearance has negative impact on sales in the mid and the tail of the sales distribution. In the "short head", a book's TV-appearance positively impacts sales. The fact whether a book was published by one of the 50 biggest publishing houses has positive impact on sales as one might expect. The coefficient increases for the top quantiles which might be attributed to differences in marketing budgets. It is natural to assume that potential top sellers receive higher budgets than niche books. Being awarded with a literary prize has a positive impact in higher quantiles, indicating that it increases sales dispersion among more popular books.

7. Discussion

Exploring the long tail phenomenon and its implications for online retailing, we have analyzed whether online reviews, discussion forums, and product recommendations help to reduce search costs and actually foster a change in the sales distribution in online book retailing. We have collected a data set containing 268,872 observations for 33,609 different books at Amazon.de. By adopting an innovative approach we were the first to develop a long tail conversion model for the German online market, based on publicly available sales data.

In order to clarify whether the proliferation of eWOM and recommendation instruments in online commerce in fact can support a paradigm change in retailing – from the steeply distributed sales of a Pareto world to a rather flattened sales distribution of a long tail world – the impact of these instruments on the unconditional sales distribution was examined. This method allows for answering questions as simple as "what are the distributional effects on sales of adding one review to each book?" or "what is the impact of increasing the proportion of books with reading excerpts?".

In sum our results indicate that eWOM actually fosters sales in the long tail, which underlines the importance of these instruments as a complement of a long tail strategy. Furthermore, with respect to our research questions, we actually see opposing effects of eWOM on the distribution of sales. On the one hand, quality information (i.e. reviews and reading excerpts) is highly valued in the long tail. Specifically, positive reviews exert strong impact in the long tail whereas negative ones do not suppress sales. This might indicate some kind of selective perception of reviews by customers. In the niche, where books are mostly found via more specific queries, critical reviews loose their bite whereas positive ones are welcome confirmations of quality. Conforming to long tail theory, an increase of the proportion of reviewed books should reduce sales inequality in the bottom of the distribution. Yet, the results also indicate that sales inequality increases in the middle of the sales distribution.

However, advertising campaigns and related offline promotion in the hit segment ($\tau = 90$) seem to create high awareness and furthermore offer customers a large information source, thereby

lessening the influence of reviews as buying persuaders. The drop in R^2 (0.26) furthermore indicates that unobserved effects, e.g. individual marketing campaigns, gain impact on sales in the top quantile.

On the other hand, popularity indicators such as review volume and discussion posts are highly influential buying arguments in the front of the sales distribution. The impacts of these variables increase in strength in the high quantiles implying an increase in sales concentration – especially among super hits. This result might have implications for further understanding the finding of Elberse and Oberholzer-Gee (2006), who discovered that the number of titles in the top ten percent of weekly VHS and DVD sales dropped by more than 50 percent from years 2000 to 2005. In other words: In the hit segment, success is concentrated in ever fewer best-selling titles (Elberse 2008) and this effect might partially be reinforced by eWOM.

In the long tail, TV appearance negatively impacts sales. A possible explanation for this fact might be a selection bias in the TV shows. There is anecdotal evidence that TV critics systematically review “arty” books of young and unknown authors which typically target a relatively small segment of customers (Clement et al., 2006). Nevertheless, most likely superstars can not be totally avoided and those books seem to profit from the promotion effect as the positive coefficients in the top quantiles suggest. However, since the data do not provide for information on the valences of the TV reviews other explanations are conceivable.

A question remains: Will online retailing totally change the economics of cultural markets and should retailers switch over from hit strategies to a re-alignment of strategic ends to the long tail phenomenon? Subsuming our results, we conclude that it seems likely that bestsellers remain important, indicating that a complete paradigm change, i.e. customers turning away from mass appeal products towards individually tailored niche products, will not occur in the end. Rather, niche products can be an additional but highly profitable revenue source which can be tapped by means of eWOM instruments.

8. References

Akerlof, G.A. (1970). The Market for "Lemons": Quality Uncertainty and the Market Mechanism. *The Quarterly Journal of Economics,* 84: 488-500.

Anderson, C. (2006). *The Long Tail.* New York, NY: Hyperion.

Ansari, A., Essegaier, S. and Kohli, R. (2000). Internet Recommendation Systems. *Journal of Marketing Research*, 37: 363-375.

Bakos, Y. (1998). The Emerging Role of Electronic Marketplaces. *Communications of the ACM*, 41: 35-42.

Bianco, A. (1997). Virtual bookstores start to get real. *Business Week,* (October 27).

Borenstein, S. and Saloner, G. (2001). Economics and Electronic Commerce. *Journal of Economic Perspectives,* 15: 3-12.

Boyd, D.E. and Bahn, K.D. (2009). When Do Large Product Assortments Benefit Customers? An Information Processing Perspective. *Journal of Retailing*, 85: 288-297.

Broniarczyk, S.M., Hoyer, W.D. and McAllister, L. (1998). Consumers' Perceptions of the Assortment Offered in a Grocery Category: The Impact of Item Reduction. *Journal of Marketing Research*, 35: 166-176.

Brynjolfsson, E., Hu, Y.J. and Smith, M.D. (2003). Consumer Surplus in the Digital Economy: Estimating the Value of Increased Product Variety at Online Booksellers. *Management Science,* 49: 1580-1596.

Brynjolfsson, E., Hu, Y.J. and Simester, D. (2006). Goodbye Pareto Principle, Hello Long Tail: The Effect of Search Costs on the Concentration of Product Sales. Working Paper, MIT Sloan School of Management, Cambridge, MA.

Chen, P. Y., Wu, S.Y. and Yoon, J. (2004). The Impact of Online Recommendations and Consumer Feedback on Sales. *Proceedings of the International Conference on Information Systems (ICIS 2004)*, Washington, D.C.

Chernev, A. (2003). When More Is Less and Less Is More: The Role of Ideal Point Availability and Assortment in Consumer Choice. *Journal of Consumer Research*, 30: 170-183.

Chernev, A. (2006). Decision Focus and Consumer Choice among Assortments. *Journal of Consumer Research*, 33: 50-59.

Chevalier, J.A. and Goolsbee, A. (2003a). Measuring Prices and Price Competition Online: Amazon.com and BarnesandNoble.com. *Quantitative Marketing and Economics*, 1: 203-222.

Chevalier, J.A. and Goolsbee, A. (2003b). Valuing Internet Retailers: Amazon and Barnes and Noble. In Baye, M.R. (Ed.) *Advances in Applied Microeconomics: Organizing the New Industrial Economy Vol. 12*. Amsterdam: Elsevier Science.

Chevalier, J.A. and Mayzlin, D. (2006). The Effect of Word of Mouth on Sales: Online Book Reviews. *Journal of Marketing Research*, 43: 345-354.

Clement, M., Proppe, D. and Sambeth, F. (2006). Der Einfluss von Meinungsführern auf den Erfolg von hedonischen Produkten. *Zeitschrift für Betriebswirtschaft*, 76: 797-824.

Dellarocas, C., Awad, N.F. and Zhang, X. (2004). Exploring the Value of Online Reviews to Organizations: Implications for Revenue Forecasting and Planning. *Proceedings of the 24th International Conference on Information Systems*, Washington, D.C.

Duan, W., Gu, B. and Whinston, A.B. (2008). The Dynamics of Online Word-of-Mouth and Product Sales – An Empirical Investigation of the Movie Industry. *Journal of Retailing*, 84: 233-242.

Elberse, A. and Oberholzer-Gee, F. (2006). Superstars and Underdogs: An Examination of the Long Tail Phenomenon in Video Sales. Working Paper, Harvard Business School, Boston, MA.

Firpo, S., Fortin, N. and Lemieux, T. (2007). Decomposing Wage Distributions using Recentered Influence Function Regressions. University of British Columbia, Vancouver, BC.

Firpo, S., Fortin, N. and Lemieux, T. (2009). Unconditional Quantile Regressions. *Econometrica*, 77: 953-973.

Greco, A.N. (1997). *The Book Publishing Industry*. Boston: Allyn and Bacon.

Hair, J. F., Anderson, R.E., Tatham, R.L. and Black, W.C. (1998). *Multivariate Data Analysis*, *fifth ed*. Upper Saddle River, NJ: Prentice-Hall.

Häubl, G. and Trifts, V. (2000). Consumer Decision-Making in Online Shopping Environments: The Effects of Interactive Decision Aids. *Marketing Science*, 19: 14-21.

Hervas-Drane, A. (2007). Word of Mouth and Recommender Systems: A Theory of the Long Tail. Working Paper, NET Institute.

Hirschman, E.C. and Holbrook, M.B. (1982). Hedonic Consumption: Emerging Concepts, Methods, and Propositions. *Journal of Marketing*, 46: 92-101.

Hu, N., Pavou, P.A. and Zhang, J. (2007). Why do Online Product Reviews have a J-Shaped Distribution? Overcoming Biases in Online Word-of-Mouth Communication. Working Paper, Singapore Management University.

Kahn, B.E. and Lehmann, D.R. (1991). Modeling Choice Among Assortments. *Journal of Retailing*, 67: 274-299.

Kawasaki, G. (2006). blog.guykawasaki. (Accessed May 6, 2008), [available at http://blog.guykawasaki.com].

Koenker, R. and Hallock, K.F. (2001). Quantile Regression – An Introduction. *Journal of Economic Perspectives,* 15: 143-156.

Li, X. and Hitt, L.M. (2004). Self Selection and Information Role of Online Product Reviews. Workshop on Information Systems and Economics (WISE 2004), Washington, D.C.

Liu, Y. (2006). Word-of-Mouth for Movies: Its Dynamics and Impact on Box Office Revenue. *Journal of Marketing*, 70: 74-89.

Lyster, M. (1999). Printed Words Get Redefined in Digital Age. *Investor's Business Daily,* (June 2) A6.

Oestreicher-Singer, G. and Sundararajan, A. (2006). Recommendation Networks and the Long Tail of Electronic Commerce. Working Paper, Tel Aviv University.

Rosenthal, M. (2008). Amazon Sales Ranks. (Accessed May 6, 2008), [available at http://www.fonerbooks.com/surfing.htm]

Schnapp, M. and Allwine, T. (2001). Mining of Book Data from Amazon.com. Paper presented at the UCB/SIMS Web mining conference. (Accessed March 24, 2006), [available at

http://www.sims.berkeley.edu:8000/resources/affiliates/workshops/webmining/Slides/ORA.ppt]

Sen, S. and Lermann, D. (2007). Why Are You Telling Me This? An Examination into Negative Consumer Reviews on the Web. *Journal of Interactive Marketing,* 21: 76-94.

Senecal, S. and Nantel, J. (2004). The Influence of Online Product Recommendations on Consumers' Online Choices. *Journal of Retailing*, 80: 159-169.

Stigliz, J.E. (1989). Imperfect Information in the Product Market. In Schmalensee, R. and Willig, R.D. (Eds.) *Handbook of Industrial Organization.* Amsterdam: Elsevier.

Stigler, G.J. and Becker, G.S. (1977). De Gustibus Non Est Disputandum. *American Economic Review*, 67: 76-90.

van Herpen, E. and Pieters, R. (2007). Anticipated Identification Costs: Improving Assortment Evaluation by Diagnostic Attributes. *International Journal of Research in Marketing*, 24: 77-88.

Welt Online (2008). Die Bestseller Methoden der Charlotte Roche. (Accessed May 3, 2008), [available at http://www.welt.de/kultur/article1888908/Die_Bestseller_Methoden_der_ Charlotte_Roche.html].

IV. BUYING WITHOUT USING: BIASES OF GERMAN BAHNCARD BUYERS*

1. Abstract

A large data set of German railway travelers is used to analyze the purchasing decision for fare-reducing BahnCards. It is expected that this tariff choice is neither completely rational nor irrational, but bounded-rational in a meaningful way. Actually the study predicts a flat-rate bias, i.e. an under-use of their BahnCards by many customers. However, this bias should not be too large and falling over time. The empirical results approve the hypotheses for the most part, especially for the more expensive BahnCard50, whereas the under-use of the cheaper BahnCard25 is so extensive that it is not worthwhile on average.

* This chapter originates from joint work with Thomas Ehrmann and Alexander Dilger.

2. Introduction

"Rules of thumb are among the more efficient pieces of equipment of optimal decision making."
Baumol/Quandt

Many firms offer consumers a menu of contracts. Phone users can choose combinations of monthly airtime minutes and prices. Households can opt for different two-part tariffs for electricity. Also railway passengers are often confronted with contractual choices concerning tariffs. They can choose between a high fixed payment ex ante accompanied by a small variable price for every mile traveled on the one hand and a "full" price for every mile traveled with no ex ante fixed fee on the other hand. The consumer has to find the scheme that minimizes the price for every mile traveled.

Given the German BahnCard scheme there is an optimal contractual choice for every (expected) quantity of miles traveled by train. The consumer's problem is to pick the right contract under conditions of imperfect information about his or her own future travel needs. A standard assumption in the economics literature is that consumers have rational expectations about their future consumption frequency and choose the utility-maximizing contract. What would be the marginal condition for the optimal decision under imperfect information? It is to equate the marginal costs of additional information gathering with its (expected) marginal benefits and then to decide accordingly. Unfortunately, this condition, that is easily formulated, is very abstract and of little help to estimate accurately the future individual travel demand. This chapter's study will check how rational the expectations are that consumers have of their future (rail) travel frequency and whether they choose their utility-maximizing contract accordingly. Moreover, it is analyzed whether their decisions have been right ex post and, if not, whether they adjust their behavior or not.

By now, scholars gained more insight into critical determinants of demand. Specifically, factors such as sunk costs are known to influence consumption as laid out by Arkes and Blumer (1985)

and McAfee, Mialon and Mialon (2010).[41] Moreover, from a behavioral economics perspective, prior payment mechanisms, bundling, and timing of payment or purchasing have been analyzed with respect to their effects on demand (Gourville and Soman, 1998; Soman, 2001; Soman and Gourville, 2001; Thaler, 1980; Thaler, 1985; Wertenbroch, 1998). For instance, studying public transport service acceptance, FitzRoy and Smith (1999) find positive impact of fixed fee season tickets on aggregate demand.[42] However, still not much is known about what affects the choice of a tariff besides expected consumption. Here we try to find out more about the effects of flat-rate pricing (Miravete, 2003). DellaVigna and Malmendier (2004; 2006) famously explain flat-rate biases, i.e. an under-use of contracts by many customers, with overconfidence about time inconsistency, using the example of memberships in fitness clubs.

Consumers may overestimate their demand for a good, e.g. due to producer advertising as argued by Mitchell and Vogelsang (1991). Drawing on empirical evidence, Nunes (2000) explains how users integrate usage expectation into the decision process when choosing between paying a flat fee for unlimited access and paying per use. Consumers tend to compare the subjective likelihood of using more than the break-even volume with the subjective likelihood of using less. He finds that consumers habitually overestimate the likelihood of using enough to justify the flat-rate and thus falsely favor this payment plan. The perceived range of usage thereby strongly affects the consumers' misperceptions.

Older studies of tariff-choice biases examined the usage of telephone services (Kling and van der Ploeg, 1990; Kridel, Lehman and Weisman, 1993; Train, Ben-Akiva and Atherton, 1989; Train, McFadden and Ben-Akiva, 1987). Lambrecht and Skiera (2006) distinguish four different causes for flat-rate biases: the "insurance effect" (Train, 1991; Miravete, 2002; Winer, 2005), the "taxi meter effect" (Prelec and Loewenstein, 1998; Thaler, 1999), the "convenience effect"

[41] Arkes and Blumer (1985) put forth that reputational concerns can justify the consideration of sunk costs in consumption decisions (on a more recent examination of social interactions see e.g. Cohen-Cole and Zanella, 2009).

[42] Public transport acceptance and pricing has widely been analyzed in the context of rising external costs of private transport due to traffic congestion (FitzRoy and Smith, 1999; see also Steimez (2009) on congestion related externalities). In a related context, Van Vuuren and Rietveld (2002) estimate the price elasticity of demand for train kilometers considering a two-part tariff option. For a general theoretical justification of differentiated price structures in the transport sector see e.g. Carbajo (1998).

(Winer, 2005; Kling and van der Ploeg, 1990), and the "overestimation effect" (DellaVigna and Malmendier, 2006). This study takes a more general look at the decision between different choices of fixed and variable fees. We expect to find a flat-rate bias, whatever its reason, but also some bounded rationality and learning, meaning better decisions over time.

This study is the first that analyzes a large data set of German railway travelers. This application differs from the so far analyzed telephone and internet services or gyms in at least two respects. First, demand seems to be more of an exogenous nature than e.g. the visits paid to a gym or the usage of internet services. Second, the monetary outlays, cancellation costs etc. are significantly higher than the respective costs of internet services. While the analysis concentrates on the cost side, some of the results also shed some light on the perceived future benefits of the BahnCard scheme.

The remainder of the chapter is organized as follows. In section 3 the simple economics of contractual choice are developed – both at the purchase date of the BahnCard and at possible renewal dates – and hypotheses are formulated. In section 4 the main features of the BahnCard data set are introduced. Then in section 5 the hypotheses are tested empirically. Section 6 discusses possible explanations for the empirical results and concludes.

3. Simple Standard Economics and Hypotheses

We set up a simple framework of contract choice and BahnCard usage. We begin with an analysis of the optimal customer's choice under complete information or ex post, given the train journeys he or she actually made.[43] Certainly, ex ante the customer does not have complete information and (not) buying a BahnCard is a risky decision. Therefore, the first derived Hypothesis 0 concerning the optimal decision under complete information is only a reference case and we do not expect this hypothesis to be true. Otherwise the analysis could stop there.

A contract (L_i, αp) gives customers the right to use a train for a fee αp, once the flat fee L_i is paid. L_i stands for different fixed fees that induce different variable fee rebates α on p. The two extreme cases are the flat-rate (L, 0), the so-called BahnCard100 (BC100) and the pay-per-mile-tariff (0, p). The most common BahnCards have either rebates α of 25% or 50% on p. Consumers can choose between all these contracts.

The discount effect of the BahnCard25 (BC25) contract with a rebate α of 25% on p begins at a minimum sum of four times the flat fee in ticket purchases. In other words, the BC25 is not worth buying for non-frequent travelers who spent less than this amount a year on rail travel. Let v be the total amount spend on rail travel a year (based on the standard fare), then the lower optimality boundary of the BC25 contract is given by

$$(1) \qquad v \geq L_{25} + 0.75v,$$

and thus,

$$(2) \qquad v_{25}^{l} = L_{25}/0.25.$$

[43] Both are not the same since buying a BahnCard changes the marginal prices. Thus complete information implies optimal decisions concerning buying a BahnCard and train tickets later, whereas the reverse is not necessarily true. One can make extra journeys with a BahnCard at hand, even if one would not have bought the card to make these journeys. However, this possible bias in this analysis does not seem to be very important empirically, because most BahnCard holders do use their cards less than optimal although the marginal prices are lower (see below).

Hence, assuming a flat fee L_{25} of EUR 57,[44] a yearly spending of EUR 228 marks the break even travel volume v_{25}^{l} for the BC25.

More frequent travelers might profit from a higher contract such as the BahnCard50 (BC50) which grants a rebate α of 50% on p. One might assume that a BC50 is worth buying if v exceeds twice the flat fee L_{50} since then total travel spending is minimized relative to the pay-per-mile tariff. However, in defining the lower optimality boundary of the BC50 contract the BC25 contract is the relevant benchmark. Consequently, the lower optimality boundary of the BC50 contract is given by

$$L_{25} + 0.75v \geq L_{50} + 0.5v, \tag{3}$$

and thus,

$$v_{50}^{l} = (L_{50} - L_{25})/0.25. \tag{4}$$

Assuming a flat fee L_{50} of EUR 225, a yearly spending of EUR 672 marks the break even travel volume v_{50}^{l} of the BC50 contract.

As can easily be seen from equation (5) and (6), only very frequent travelers profit from a BC100 contract. The lower optimality boundary is given by

$$L_{50} + 0.5v \geq L_{100}, \tag{5}$$

and thus,

$$v_{100}^{l} = (L_{100} - L_{50})/0.5. \tag{6}$$

At a price of EUR 3,650, this card is only worth buying if v exceeds EUR 6,850.[45]

To subsume: Based on current prices, for every v smaller than EUR 228 the pay-per-mile tariff is optimal, for v between EUR 228 and EUR 672 the BC25 contract is optimal, and for v in the

[44] This is the current (2009) value. In the past the BC25 as well as the other BahnCards were cheaper but the formulas remain the same.
[45] The optimality intervals of equations (1) to (6) apply for all $L_{50} \geq 2 \cdot L_{25}$ and $L_{100} \geq 2 \cdot L_{50}$.

interval of EUR 672 and EUR 6,850 the BC50 is optimal. For every *v* beyond EUR 6,850 the BC100 minimizes travel costs.

Hypothesis 0 (Optimal Buying Decision for a Particular BahnCard): Agents choose the optimal BahnCard contract for the mileage they travel by train.

It shall be stressed again that this study does not expect that all or most customers decide optimally ex post. Hypothesis 0 is the reference case and we expect empirical evidence contrary to it. Nevertheless, the lack of complete information and the existence of bounded rationality do not mean the complete lack of all information and total irrationality. On the contrary, we expect some (bounded) rationality even in the mistakes and biases human beings are prone to.

That is why we expect that BahnCards and travel demand are systematically connected. Who anticipates more miles *m* by train will buy with a higher probability a more expensive BahnCard than someone expecting to travel less miles. Moreover, whereas the individual expectations can be wrong, in the aggregate they are fulfilled more or less. There is a second reason why the owners of more expensive BahnCards will use them more than those of cheaper ones: The marginal price of an additional mile is lower. Unfortunately, we cannot differentiate empirically between these two explanations but both support the following hypothesis.

Hypothesis 1 (Higher Travel Demand by Users of More Expensive BahnCards): The different contracts (L_i , αp) *require different degrees of ex ante commitment of consumers and change marginal prices such that the buyers of more expensive BahnCards accordingly travel more.*

$$m_{100} > m_{50} > m_{25} > m_0. \tag{7}$$

Whereas Hypothesis 1 implies at least a weak form of rationality, we expect some deviations from strict rationality, first and foremost a flat-rate bias. A flat-rate bias means that many travelers prefer a (more or less) flat rate even though their billing rate would be lower on a pay-per-mile price.[46] The magnitude *F* of a flat-rate bias can be measured by the additional price of the

[46] According to Nunes (2000) a flat-rate bias exists on the individual level when an actor pays a fixed fee for unlimited access that costs more than a usage-dependent tariff would have cost, given his demonstrat-

chosen BahnCard and bought tickets compared to the optimal BahnCard and corresponding ticket prices. For example, buying once a BC100 ends any worries about the costs of all train trips in the following 12 months. This ease of mind has some value such that paying more than the savings per miles traveled can be boundedly rational.[47] In the case of the BC100 it is also possible to save real transaction and opportunity costs because its owner does not need to buy any more tickets. BC50 and BC25 do not bring about this real convenience but also lower worries and emotional costs for many train tickets.

Hypothesis 2 (Existence of a Flat-Rate Bias): There exists a flat-rate bias among BahnCard holders, i.e. there are many BahnCards which are more expensive than optimal (considering only the BahnCard and ticket fees).

Moreover, we expect that the flat-rate bias is not independent of the kind of BahnCard. Instead F should be higher for more expensive BahnCards. First, it is possible to make greater mistakes with a more expensive card. In the extreme, a BahnCard is not used at all such that the flat-rate bias equals the price of the card. Second, there is a further possibility of mistake when buying a BC50 compared to a BC25. As shown above, it is possible that a BC50 is better than no BahnCard at all but worse than buying a BC25. The same holds for even a larger range of miles when a BC100 is compared to a BC50.

Hypothesis 3 (Higher Flat-Rate Bias of More Expensive BahnCards): The flat-rate bias is increasing in the price of the BahnCard.

(8) $F_{100} > F_{50} > F_{25}$.

ed demand. Here we follow Nunes (2000) and extend this definition to our context. Hence, a flat-rate bias exists when an actor chooses a contract involving a high fixed fee and low variable costs although another contract with a lower fixed fee and higher variable costs would have resulted in a lower billing rate, given his demonstrated demand. In measuring the existence of a flat-rate bias, we basically follow the common method of measuring the proportion of users in a tariff that would have paid less in a lower tariff, given ex post usage data (see e.g. Mitchell and Vogelsang 1991; Kridel, Lehman and Weisman, 1993; Nunes, 2000; DellaVigna and Malmendier, 2006).

[47] Perfect rationality includes the ability to calculate everything and to make decisions without costs, emotional or otherwise (besides the real costs one is deciding about, of course).

Whereas we expect a flat-rate bias and under-using of many BahnCards, we do not think that the mistakes of BahnCard users are arbitrarily high. On the contrary, the average user probably profits not only from having a BahnCard (compared to having none) but also from having the specific BahnCard he or she has (compared to none or any other). This also means that those customers, who have the right BahnCard profit more by it than the others, having the wrong one, lose. From this it follows that the average utility from a BahnCard is positive.

Hypothesis 4 (BahnCards are Useful on Average): The average utility for BahnCard users is higher with their specific BahnCard than without it.

What can be said about the different customers' expectations about their own future travel demand? Customers who expect a low travel demand in the future prefer a cheaper BahnCard or none at all. Therefore the option to switch sooner to a lower BahnCard or to payment per mile should be valued highly by them. Customers who view themselves as heavy users in the future should prefer the high BahnCard contract types. These customers should value the reduced price of each mile traveled and should not mind the yearly commitment. A kind of sorting therefore implies that the higher the chosen contract in the past the more likely these customers should renew a high contract. The renewal probability R of a high contract should be higher than for the low contract.

Hypothesis 5 (Higher Renewal Probabilities for More Expensive BahnCards): The renewal probability of a more expensive BahnCard is higher than for a cheaper one.

$$R_{100} > R_{50} > R_{25}. \tag{9}$$

There are other questions about what is happening at the end of the contractual period, after every customer knows at least for the past about his or her travel demand. Each consumer can then either renew the BahnCard with the same or a different contract, or opt out (i.e. switch to the pay-per-mile contract). Let s_i denote the annual savings from switching from contract i to another contract. Then the customer should switch if the value of savings is positive, i.e.

$$s_i > 0. \tag{10}$$

Should the customer incur switching costs k than he or she has to compare realized k with the benefits, implying the switching becomes more costly. Now over time even bounded-rational holders of BahnCards should learn something about their usage and should learn to make more and more informed guesses about s_i and therefore about their optimal contract. Learning will lead customers with ex post low usage to opt for $(0, p)$. When we define individuals that hold on to their initial BahnCard contract as "constant customers" than we should find in the cohort of a particular BahnCard higher usage in subsequent years than was the usage of the initial group, because many of the less frequent travelers switch to pay per mile over the years. Among BahnCard users that have initially opted for an expensive contract, the expected usage in subsequent years among "constant customers" is higher than the usage in the first year for the initial group.

Hypothesis 6 (Better Usage of BahnCards Over Time): In the aggregate the use of the BahnCards gets better over time (time measured in calendar years or the years of using a particular type of BahnCard).

Given that someone holds his or her BahnCard for two or more periods, we expect fewer mistakes in using it in the long run than within single periods. The estimation of one's travel demand can be too high or too low but in the long run such mistakes cancel each other out. Moreover, bounded-rational people are able to learn from their mistakes such that their use of a BahnCard gets better over time.

Hypothesis 7 (Better Use Over Longer Periods): For individuals who hold their BahnCards, the use is better over longer periods of time than in single periods.

4. BahnCard Dataset

4.1 Data and Sample Period

Our data to test the hypotheses were provided by the German railway company Deutsche Bahn (DB) AG and comprise detailed information on customers' individual demographic characteristics, BahnCard contract choices and individual transaction data, i.e. ticket purchase behavior. The representative sample was drawn from the population of members of the company's customer loyalty programs "bahn.bonus" and "bahn.comfort". The bahn.bonus program awards points to customers based on the amount they purchase. These points can be collected and finally spend on different rewards, e.g. train tickets, 1st class upgrades or car rentals. bahn.comfort is a customer program which awards premium status for customers who spend a predetermined amount of money on ticket purchases. Within both programs every euro in purchases translates in one point.[48] Since customers are rewarded with points for ticket purchases the data set allows for the reproduction of individual traveling behavior. Based on this information we evaluate the efficiency of BahnCard contracts. The sample period is December 2002 through July 2008.

4.2 Contractual Menu

The data include contracts with different standard reductions on tickets. Specifically, customers can choose between the following contracts: BC25, BC50 and BC100. Each contract is available for the 1st and 2nd class. Currently, the standard BC25 contract involves an up-front fee of EUR 57 (2nd class) and a reduction of 25 percent on domestic tickets for 12 months from the date of issue.[49] The BC50 is offered at a price of EUR 225 (2nd class) and grants reductions of 50 percent on fares. The BC100 costs EUR 3,650 (2nd class) and allows free travel on all DB-trains for one year.

In addition to the standard contracts, the DB AG offers several reduced fee contracts with initial fees ranging from EUR 0 (BC25 2nd class for family members of BC100 customers) to EUR

[48] Within the bahn.bonus program purchases of 1st class tickets are awarded with 1.5 points.
[49] This accounts at least for all DB-trains. Parts of the regional passenger rail are operated by other rail companies than the DB AG. Several of these companies grant reductions on fares for BahnCard customers, too.

115 (BC50 2nd class for students and senior citizens or e.g. family members of BC50 customers). If not cancelled, BC25 and BC50 contracts are automatically renewed after 12 months. Cancellation can be done in written form until six weeks before the end of validity. BC100 contracts end after one year.[50] BC25 and BC50 customers can switch to higher contracts within the contract period. The residual value of the current card is then refunded. Customers cannot switch to lower contracts during the duration of an ongoing contract.

4.3 Sample Construction and Key Variables

In sum, we received data on more than four million transactions, each being related to one of approximately 800,000 BahnCards. These data constitute the travel history of over 300,000 customers. However, to construct a reliable data base for our analysis we had to make some severe adjustments. Since not all members of the loyalty programs frequently collect bonus points we excluded all customers whose overall lifetime sales volume equals zero. Furthermore, we dropped all customers with non-standard and promotional contracts such as cards with less than 12 months duration, for example. Finally, we concentrated on customers with 2nd class BahnCards, in order to avoid assignment problems and to achieve a maximum comparability between contracts. Consequently, our final dataset features 259,752 BahnCards of 83,263 customers with corresponding transaction data.[51]

With regards to single transactions, i.e. ticket purchases, key variables of the dataset include the purchase date, ticket price, reduction rate, class, origin and destination station, number of passengers, a round trip indicator, and a general description indicating whether the travel's origin or destination station is not domestic, for example. Since the ticket price represents the reduced price we calculated the standard fare based on the actual reduction rate for each ticket in order to have a common basis of comparison. Inbound and outbound tickets, i.e. travels from or to another country, constitute a special case here. The DB AG offers a "Railplus" option for Bahn-

[50] The BC100 can also be subscribed at a monthly fee of EUR 335 with minimum contract duration of 12 months.

[51] Despite our adjustments there still might be some unobserved aspects within the data. First, the data originates from a decentralized large scale system which generally aggravates consistent data input. Second, a potential limitation is that customers might exhibit inconsistencies with respect to bonus point collection behavior.

card customers, which grants a 25% reduction on standard fares for the abroad section when traveling to 29 European countries.[52] Obviously, for the calculation of standard fares of BC50 customers this poses issues. In particular, the ticket price reflects a 50% reduction on the standard fare of the domestic section of the travel and a 25% reduction on the standard fare of the abroad section. Since we do not have any information on the ratio of domestic to non-domestic route length for these tickets, we assumed that customers realized a reduction of at least 25% on the entire route.[53]

For our analysis we aggregated all transactional data on a contract basis. Consequently, our final sample features the following information for each BahnCard contract: Customer ID, reduction rate, flat fee, date of issue, end of validity, a dummy indicating whether the contract is subject to automatic renewal, a Railplus dummy, total spending on tickets (including reduction), and virtual total spending on tickets without a BahnCard (hypothetical standard fare).[54]

Our first research question is, whether the chosen contracts are optimal from the customers' perspectives considering individual traveling behavior. Hence, we calculated the sums of total spending for all possibly available contracts as benchmarks for the actual sum of spending and assigned dummy variables indicating whether the chosen contract was efficient, i.e. cost minimizing, or not.[55] First, one dummy indicates whether a lower contract implies lower costs, i.e. if a customer did not reach the break even travel volume. Second, one dummy indicates whether or not a higher contract implies lower costs, i.e. if a customer reached the break even travel volume of a higher contract. Third, a set of variables constitute the difference in costs between the actual contract and the alternative options. These variables serve as bases for utility calculation.

[52] Since December 2007 Railplus is included in standard BahnCard contracts. Before that, Railplus was optional and subject to a fee of EUR 15.

[53] This assumption tends to underestimate reductions of BC50 customers. However, the alternative option would have been to exclude all in- and outbound tickets, which, we feel, would have been an even stronger underestimation of Bahncard reductions. Anyway, the proportion and costs of such trips abroad are quite small (see below).

[54] Obviously, BC100 users do not need to purchase tickets and consequently our database lacks information on travel behavior of those customers. From this follows that we cannot draw conclusions on the individual efficiency of BC100 contracts. Nevertheless, these contracts function as a reference point within the data and represent an important alternative option for BC25 and BC50 customers at the end of a contract period.

[55] Several contracts are ongoing at the end of the sample period. In these cases, we included the flat fee on a pro rata basis.

Apart from the evaluation of single contract selection, we are especially interested in the individual contract history of customers. Hence, each dataset contains information on the successive contracts which allows for observing if customers switch inefficient contracts or opt out at the end of a contract period in order to increase utility.

4.4 Descriptive Statistics

Within our sample, 45.9 percent of all contracts with transactional data are BC25 contracts and 54.1 percent BC50. Table 8 shows the descriptive statistics for the sample.

Contracts	BC25	BC50
number of contracts		
total	119,240	140,512
completed	101,597	114,560
initial fee	48.04	140.72
	(14.86)	(50.17)
total spending on tickets		
incl. reduction	102.36	236.28
	(206.53)	(402.33)
hypothetical standard fare	136.48	467.78
	(275.38)	(797.49)
total spending on abroad tickets	4.62	7.17
	(37.53)	(58.46)
total costs	145.39	363.77
	(210.03)	(419.49)
number of tickets	2.18	7.25
	(5.08)	(12.94)
n	119,240	140,512

Notes: Standard deviation in parentheses. A completed contract ends before the date of data retrieval.

Table 8. *Descriptive Statistics Contracts*

On average, BC25 customers paid EUR 48.04 as initial fee, BC50 customers paid EUR 140.72. A typical BC25 customer spends EUR 102.36 on tickets, BC50 users spend EUR 236.28. The average spending on in- and outbound tickets is considerably small with EUR 4.62 (BC25) and EUR 7.17 (BC50). Each contract accounts for 2.18 (BC25) and 7.25 (BC50) tickets and gene-

rates total costs for the average costumer of EUR 145.39 (BC25) and EUR 363.77 (BC50), respectively.

Customers	
age at first contract	38.66 (17.59)
female	0.54 (0.50)
total spending on tickets	544.70 (1057.92)
number of contracts	3.12 (1.69)
N	83,263

Notes: Standard deviation in parentheses. Total spending on tickets includes reduction.

Table 9. *Descriptive Statistics Customers*

With respect to customers, Table 9 reveals that the typical BahnCard customer is approximately 39 years old at first sign-up. About 54% of customers are female and individuals buy on average 3.12 BahnCards during the sample period. The average total spending on tickets for all contracts is EUR 544.70.

5. Empirical Results

We used the data described in section 4 to test the hypotheses derived in section 3 and got the following results.

Hypothesis 0 (Optimal Buying Decision for a Particular BahnCard):

If BahnCards were used perfectly – i.e., from an ex post perspective, users choose the right contract for the mileage they travel by train – we should not observe sub-optimal contracts. However, our first empirical result shows that the observed distribution of optimal contracts deviates significantly from the assumption of perfect usage.[56] A chi-square goodness-of-fit test confirms that the null hypothesis, i.e., the frequency distribution of optimal contracts in our sample is consistent with the theoretical distribution (100%), can be rejected at a significance level of $p<.001$.[57] In particular, we assigned a dummy variable indicating whether a certain contract is cost minimizing (*optimal*=1) or not (*optimal*=0), depending on individual travel behavior. That is, another contract, either a higher, a lower, or the pay-per-mile contract would have resulted in lower total costs. In fact, only 29.7% of contracts were within the respective optimality boundaries as derived in section 3. Hence, 70.3% of contracts were not optimal. As stressed above, we actually did not expect that all users decide optimally ex post, and consequently, Hypothesis 0 is not supported by our data.

Hypothesis 1 (Higher Travel Demand by Users of More Expensive BahnCards):

As noted in section 4, the data unfortunately do not provide information on travel behavior of BC100 and pay-per-mile tariff users. Nevertheless, according to equation (7), buyers of a BC50 should travel more than BC25 customers, reflecting higher degrees of ex ante commitment as

[56] The methods applied in this chapter arise from the specific hypothesis formulations as well as from the scale of the criterions under consideration. Specifically, when comparing rates and proportions of the occurrence of a particular attribute or a combination of attributes, we resort to the class of chi-square methods, representing the most frequently applied statistical tests for categorical data (e.g. Fleiss et al., 2003). However, considering interval-scaled data we apply both parametric and non-parametric methods in order to enhance the robustness of our results.

[57] A chi-square "goodness-of-fit-test" tests whether the frequency distribution of certain events observed in a sample is consistent with a particular theoretical distribution. Thereby, the events have to be mutually exclusive with a total probability of 1. Since the approximation to the chi-square distribution breaks down if expected frequencies are too low (<5, see Siegel and Castellan, 1988; or Gibbons and Chakraborti, 2003), we slightly relaxed the strong assumption of no sub-optimal contracts to the lower boundary.

well as lower marginal prices. Table 10 illustrates the results of an independent-samples t-test.[58] The test variable is *total annual spending on tickets* (based on the standard fare), with BC50 and BC25 contracts representing the two samples. BC25 customers spend on average the amount of EUR 144.30 on tickets, this being significantly less than the average spending of BC50 users (EUR 484.51). Consequently, we find strong support for Hypothesis 1.

Hypothesis 1	N	Mean	s.e.	p*
annual ticket spending BC25	101,597	144.30	.898	
annual ticket spending BC50	114,560	484.51	2.449	
difference		-340.21	2.608	<.001

** p values from independent samples t-test.*

Table 10. *Empirical Results – Usage by Contracts*

Hypothesis 2 (Existence of a Flat-Rate Bias):

While customers obviously exhibit weak forms of rationality, the reference case indicates deviations from strict rationality in that users seem to be prone to under-using of BahnCards, i.e. total monetary costs are not minimized. Specifically, as it might be boundedly rational to prefer a BahnCard with lower marginal costs over the pay-per-mile price we should observe a flat-rate bias. Table 11 shows the results of a non-parametric sign test, which is commonly used when only the direction of differences is of interest, not the magnitude (Conover 1999). We constructed two variables *current BahnCard* and *optimal BahnCard* and assigned values (0, 25, 50, and 100) according to the reduction rates of the actual contract and the contract which would have minimized total (monetary) costs, depending on individual travel behavior. The sign test poses the null hypothesis *optimal BahnCard* = *current BahnCard* and tests whether positive and negative differences between a random pair of measurements are equally likely to occur. With respect to Hypothesis 2 we should observe a higher probability for the case *optimal BahnCard* < *current BahnCard.*

[58] Since we only included completed contracts, n scales down to the total of 216,157 observations. The underlying assumptions of the t-test are homogeneity of variances in the two samples and the criterion under consideration has to be interval-scaled. Beyond that, the t-test requires normal distribution of the sample-mean, which is given for moderately large sample sizes (n>30) according to the central limit theorem (Greene, 2008; Lumley et al., 2002). However, we corrected for heterogeneous variances, and we furthermore conducted a non-parametric Mann-Whitney-U-Test in order to test the robustness of our results (e.g. Daniel, 2000). The results did not change.

Hypothesis 2	n
optimal BahnCard < current BahnCard	173,969
optimal BahnCard > current BahnCard	8,643
optimal BahnCard = current BahnCard	77,140
n	259,752
z	-386.88
p	<.001

Table 11. *Empirical Results – Sign Test*

In 67% of cases customers under-used their BahnCard, i.e., switching to a lower contract or to the pay-per-mile tariff would lower total monetary costs. In only 3.3% of cases BahnCards were over-used, i.e. switching to a higher contract would minimize costs. As noted above, 29.7% of contracts are optimal. The null hypothesis can be rejected at a significance level of $p<.001$. Thus, our data corroborate the existence of a flat-rate bias and consequently Hypothesis 2.[59]

Hypothesis 3 (Higher Flat-Rate Bias of More Expensive BahnCards):

In order to test Hypothesis 3, which states that the flat-rate bias is higher for more expensive BahnCards, we analyzed whether the magnitude of the bias *F* differs significantly between BahnCard types. Therefore we conducted an independent-samples t-test with the test variable being defined as the difference in costs between the optimal contract plus corresponding tickets purchases and the chosen BahnCard and bought tickets. As we only consider under-used contracts *n* scales down to 173,969 with BC25 (88,828) and BC50 (85,141) contracts representing the two samples. As can easily be seen from Table 12, our results support Hypothesis 3 ($p<.001$). On average, BC25 users with under-used contracts incurred losses of EUR 35.45. BC50 users incurred losses of EUR 96.68 as a result of choosing sub-optimal contracts (from a strictly monetary perspective).[60] This appears reasonable since the higher flat-fee allows for more severe mistakes from an individual perspective.

[59] Furthermore, we conducted another chi-square goodness-of-fit test which confirms that the null hypothesis, i.e., the frequency distribution of sub-optimal contracts in our sample is consistent with the theoretical distribution (0%), can be rejected at a significance level of $p<.001$. The test variable was a dummy indicating whether a contract was under-used (*under-use*=1) or not (*under-use*=0).

[60] A further Mann-Whitney-U-Test was highly significant, too.

Hypothesis 3	N	Mean	s.e.	p*
utility BC25 under-use	88,828	-35.45	.064	
utility BC50 under-use	85,141	-96.68	.192	
difference		61.23	.202	<.001

** p values from independent samples t-test.*

Table 12. *Empirical Results – Flat-Rate Bias by Contracts*

Hypothesis 4 (BahnCards are Useful on Average):

Although we observe a flat-rate bias, we expect the average user to profit not only from having a BahnCard at all but also from having his or her specific BahnCard. If so, users with the right BahnCard should profit more by it than others lose, as predicted by Hypothesis 4.

For a start, we analyze whether customers profit from having a BahnCard at all, i.e., as compared to having none. We conducted an independent-samples t-test with users who under-used their BahnCards (*under-use*=1) and those who did not (*under-use*=0) representing the two samples. The test variable is defined as the absolute value of the difference in costs of the chosen BahnCard plus ticket prices and the costs of the ticket purchases based on the standard fare.[61] Considering both BC25 and BC50 contracts, the empirical results in Table 13 support Hypothesis 4. On average, customers who under-use their BahnCard incur losses of EUR 65.69, whereas customers with the right contract benefit from it to the amount of EUR 272.15 (p<.001).[62] In the aggregate, the mean of losses and profits is positive, i.e., the average utility from a BahnCard amounts to EUR 52.17. However, splitting the sample into sub-samples according to contract type, we obtain the following results: within the BC25 sample losses due to under-usage on average amount to EUR 35.45, profits to EUR 68.61. Although losses and profits differ significantly (p<.001), Hypothesis 4 is only partly corroborated with respect to this sub-sample since its aggregate mean is negative with EUR -8.91. Within the BC50 sample losses are on average EUR 99.15, profits amount to EUR 374.95 (p<.001). Here the aggregate mean is positive again with EUR 104.01.

[61] Note that within these tests we do not evaluate over-usage as faulty behavior.
[62] A Mann-Whitney-U-Test yielded the same result.

Hypothesis 4	N	Mean	s.e.	p*
Point of Comparison: Standard Fare				
utility (absolute value) under-use	169,128	65.69	.126	
utility (absolute value) optimal use	90,624	272.15	1.358	
difference		-206.45	1.363	<.001
utility (absolute value) BC25 under-use	88,828	35.45	.064	
utility (absolute value) BC25 optimal use	30,412	68.61	.554	
difference		-33.16	.557	<.001
utility (absolute value) BC50 under-use	80,300	99.15	.198	
utility (absolute value) BC50 optimal use	60,212	374.95	1.889	
difference		-275.80	1.899	<.001
Point of Comparison: Lower Boundary				
utility (absolute value) under-use	173,969	49.29	.090	
utility (absolute value) optimal use	85,783	142.58	.681	
difference		-93.29	.687	<.001
utility (absolute value) BC25 under-use	88,828	35.45	.064	
utility (absolute value) BC25 optimal use	30,412	68.61	.554	
difference		-33.16	.557	<.001
utility (absolute value) BC50 under-use	85,141	63.73	.158	
utility (absolute value) BC50 optimal use	55,371	183.21	.968	
difference		-119.48	.981	<.001
Point of Comparison: Optimal Contract				
utility (absolute value) under-use	173,969	65.41	.124	
utility (absolute value) optimal use	85,783	142.58	.681	
difference		-77.17	.692	<.001
utility (absolute value) BC25 under-use	88,828	35.45	.064	
utility (absolute value) BC25 optimal use	30,412	68.61	.554	
difference		-33.16	.557	<.001
utility (absolute value) BC50 under-use	85,141	96.68	.192	
utility (absolute value) BC50 optimal use	55,371	183.21	.968	
difference		-86.53	.987	<.001

** p values from independent samples t-test.*

Table 13. *Empirical Results – Average Utility*

Now we analyze whether the average customer profits from having his or her specific BahnCard. Defining the relevant point of comparison is crucial for quantifying the (dis-)utility derived from a certain contract. This is all the more true as customers differ in terms of utility assessment, either due to differences in search costs or the level of rationality for example. Hence, from the diversity of individual perceptions follows a variety of potential benchmarks for measuring utility. In the following we account for this fact by, firstly, comparing the actual contract to the next lower contract, and, secondly, by comparing the actual contract to the ex post optimal regime.

Within this test, the test variable is defined as the absolute value of the difference in costs of the chosen BahnCard and ticket purchases and the price of the next lower contract and corresponding tickets. Considering both BC25 and BC50 contracts, the empirical results in Table 13 support Hypothesis 4. Users with the wrong BahnCard on average incur losses of EUR 49.29, whereas customers with the right contract benefit from it to the amount of EUR 142.58 ($p<.001$).[63] In the aggregate, the average utility from a BahnCard is positive and amounts to EUR 14.08. Further differentiating between contracts, we obtain the following results: within the sample of the BC25 users, results do not change as compared to the first test.[64] Those with the wrong contract lose EUR 35.45, this being significantly less than the positive utility of optimal users which is EUR 68.61 ($p<.001$). The aggregate mean is negative with EUR -8.91. Within the BC50 sample, the aggregate mean is positive again (EUR 33.58) and the positive utility of travelers with the right contract (EUR 183.21) exceeds the negative of those with the wrong contract (EUR 63.73) significantly ($p<.001$).

Next, we calculate the utility of the actual contract based on the comparison to the ex post optimal regime. In particular, for under-using BC50 customers either the BC25 or no BahnCard is the cost minimizing option. Consequently, the test variable is defined as absolute value of the difference in costs of the chosen BahnCard and ticket purchases and the price of the ex post optimal contract and corresponding tickets.[65] As Table 13 shows, the results basically remain stable. Specifically, losses due to under-usage rise to an average of EUR 65.41, this being significantly less than profits which remain constant as compared to the foregoing test with EUR 142.58 ($p<.001$). In the aggregate users still profit from their specific BahnCard as the average utility is positive with EUR 3.27. This decrease in average utility is driven by the rising average losses due to under-usage of the BC50 sub-sample, since the results of the BC25 contracts do not change. In particular, losses are on average EUR 96.68 as opposed to profits of EUR 183.21

[63] A Mann-Whitney-U-Test yielded the same result.
[64] Note that for a BC25 contract the next lower boundary is the no BahnCard option, i.e. the standard fare.
[65] For optimally used contracts, utility is calculated based on the comparison to the next worst contract, i.e. the lower boundary.

(p<.001). As a consequence the average utility of the BC50 sample is positive and amounts to EUR 13.62.

To summarize, Hypothesis 4 is mainly corroborated by our data, i.e., the average user not only profits from having a BahnCard at all but also from having his or her specific BahnCard. This applies with the exception that the average utility of the BC25 sub-sample is negative.

Hypothesis 5 (Higher Renewal Probabilities for More Expensive BahnCards):

To test whether more expensive contracts imply higher renewal probabilities, as stated by Hypothesis 5, we carried out a chi-square test of independence. We assigned a dummy variable indicating whether a customer holds on to his or her initial contract (*renewal*=1) or not, i.e., whether the user opts out or switches the contract (*renewal*=0). Contract type (25, 50) was the second variable under consideration in our contingency table. Hypothesis 5 is substantiated by the data. The renewal probability increases for more expensive BahnCards. In particular, BC50 customers are 1.1 times as likely to renew contracts as BC25 customers (p<.001, see Table 14).

Hypothesis 5		renewal	no	yes	total
contract	BC25	n	11,498	28,004	39,502
		% of contract	29.1	70.9	100
	BC50	n	8,351	30,026	38,377
		% of contract	21.8	78.2	100
total		n	19,849	58,030	77,879
		% of contract	25.5	74.5	100
χ^2			553.27		
df			1		
p			<.001		
φ			.084		

Table 14. *Empirical Results – Contract Renewal Probability*

Hypothesis 6 (Better Usage of BahnCards Over Time):

In order to analyze usage of BahnCards over time and to assess the influence of learning we carried out different tests. From a general perspective BahnCard usage should increase in calendar years if learning occurs, as stated in Hypothesis 6. In particular, we compare the share of

under-used contracts in the years 2003 and 2007.[66] We constructed a 2x2 contingency table with a dummy indicating whether a contract was under-used (*under-use*=1) or not (*under-use*=0), and *calendar year* (2003, 2007) as variables (see Table 15). The results of a chi-square test of independence support Hypothesis 6 since the null hypothesis, that there is no relationship between the row and the column frequencies, can be rejected at a significance level of $p<.001$ (φ=-.076). Specifically, the share of under-used BahnCards decreases from years 2003 (69.9%) to 2007 (62.6%) by 7.3 percentage points.

Hypothesis 6		under-use	no	yes	total
year	2003	n	12,024	27,880	39,904
		% of year	30.1	69.9	100
	2007	n	18,825	31,562	50,387
		% of year	37.4	62.6	100
total		n	30,849	59,442	90,291
		% of year	25.5	74.5	100
χ^2			517.30		
df			1		
p			<.001		
φ			-.076		

Table 15. *Empirical Results – Better Usage Over Time, Calendar Years*

However, focusing on individual experience with using a particular type of BahnCard, we obtain mixed results with respect to Hypothesis 6. As explained above, we expect higher usage in subsequent years for the cohort of a particular BahnCard as compared to the initial group since users with ex post low usage should opt out or downgrade.

Table 16 and Table 17 show the results of two one-way analysis of variance (ANOVA) tests.[67] The test variable is *total annual spending on tickets* (based on the standard fare) with first,

[66] Note that these are the first and the last completed calendar years in our database.

[67] The standard ANOVA test assumes equal variances within the samples. Yet, a Levene-test shows that the assumption of homogenous variances is violated here. Generally, the F-test reacts robust to this violation (Hair et al., 2009). However, as sample sizes differ significantly here, we applied the Welch statistic which is more powerful under these circumstances. Furthermore, the post-hoc analysis is carried out by means of the Tamhane T2-test which does not assume equal variances (Tamhane, 1977).

second and third year contracts being the three samples. The first test considers the BC25 cohort, the second the BC50 cohort.[68]

Hypothesis 6	N	Mean	s.e.	p*
annual ticket spending 1st contract	26,393	160.62	1.868	
annual ticket spending 2nd contract	17,924	119.45	1.960	
annual ticket spending 3rd contract	12,396	124.88	2.490	
post-hoc-tests				
1st vs. 2nd		41.16	2.707	<.001
1st vs. 3rd		35.74	3.113	<.001
2nd vs. 3rd		-5.429	3.169	.238
F	132.776			
Welch-Test	131.826			
df	2			
p**	<.001			

** p values from Tamhane post-hoc test.*
*** p values from oneway ANOVA Welch-statistics. p value from standard F statistic <.001.*

Table 16. *Empirical Results – Better Usage Over Time, BahnCard25 Cohort*

With respect to the BC25 cohort, it can be seen that the average annual spending on tickets of the group of users that hold on to their contract in the second year is significantly lower than that of the initial group.

Hypothesis 6	N	Mean	s.e.	p*
annual ticket spending 1st contract	23,801	438.92	4.815	
annual ticket spending 2nd contract	18,859	427.70	5.901	
annual ticket spending 3rd contract	13,789	475.79	7.261	
post-hoc-tests				
1st vs. 2nd		11.22	7.616	.366
1st vs. 3rd		-36.87	8.712	<.001
2nd vs. 3rd		-48.09	9.356	<.001
F	15.532			
Welch-Test	13.917			
df	2			
p**	<.001			

** p values from Tamhane post-hoc test.*
*** p values from oneway ANOVA Welch-statistics. p value from standard F statistic <.001.*

Table 17. *Empirical Results – Better Usage Over Time, BahnCard50 Cohort*

In particular, the mean decreases from EUR 160.62 to EUR 119.45 (p<.001). In this respect, we do not find support for Hypothesis 6. Nevertheless, we see the average spending on tickets

[68] The cohort samples were constructed as follows: First year BahnCards were included if contracted before year 2005 in order to ensure completeness of all potential successor contracts. Second and third year contracts were only included if the precedent contract was of the same type and if there was no gap between the contracts.

slightly rising from year two to three, although the difference in means fails to match common significance levels.

However, turning to the BC50 cohort we do find tentative support for Hypothesis 6. Although spending slightly decreases from year one to two (p=.366), usage increases significantly from EUR 427.70 in the second year to EUR 475.79 in the third year (p<.001).

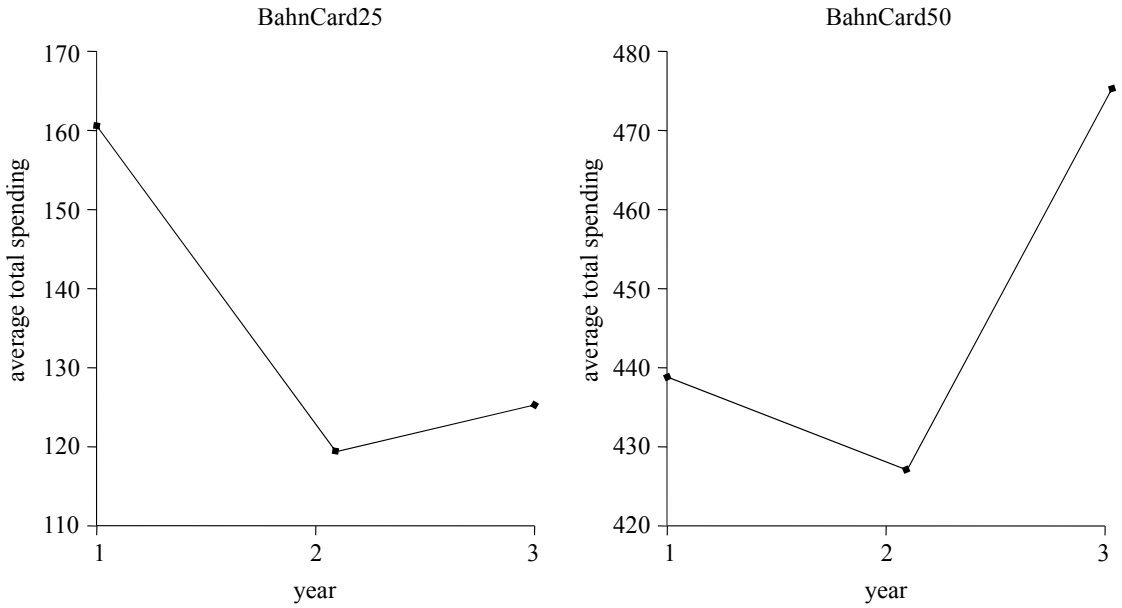

Figure 12. *Mean Diagrams Oneway ANOVA – Total Spending*

As Figure 12 illustrates there is a u-shaped development of annual spending in both cohorts. An explanation for these unexpected results could be that some users simply forget to cancel their cards whereas buying the first card needs a conscious decision. However, forgetting to cancel twice (i.e. again at the end of the second period) becomes more painful as the price of the BahnCard rises. Hence, this might be the reason why the increase in spending from the second to the third year is stronger for the BC50 cohort.

Hypothesis 7 (Better Use Over Longer Periods):

In the long run, BahnCard usage should improve from an individual perspective. On the one hand, inaccuracies in demand estimations might even out and, on the other hand, given that people learn from their mistakes, customers adapt by either switching contracts or modifying their buying behavior. A paired-samples t-test supports this (Table 18). The sample consists of

all customers who had at least three BahnCards. The test variables under consideration are as follows: The first variable is based on the difference in costs between the chosen BahnCard and ticket purchases and the price of the next lower contract and corresponding tickets, i.e., the *utility of the first BahnCard.* This variable is negative for under-used contracts and positive for optimal contracts. The second variable is the *average utility of the first three BahnCards*. As is illustrated by Table 18, the average utility of a first year BahnCard amounts to EUR 11.51. This is significantly lower than the average utility within the first three contract years which equals EUR 14.69 (p<.001).[69] Hence, for individuals who hold on to a BahnCard contract, the use is better over longer periods of time.

Hypothesis 7	N	Mean	s.e.	p*
utility 1st contract	46,021	11.51	.672	
average utility contracts 1-3	46,021	14.69	.626	
difference	46,021	-3.18	.404	<.001

* *p values from paired samples t-test.*

Table 18. *Empirical Results – Better Usage Over Longer Periods*

[69] A Wilcoxon signed-rank test (e.g. Gibbons and Chakraborti, 2003) confirms the result at a significance level of p<.001.

6. Discussion and Implications

There is empirical evidence in favor of most of our hypotheses. Most buyers of a BahnCard did not make the ex post optimal buying decision but they showed a flat-rate bias. Nevertheless, their decisions were not completely irrational but they showed, at least in the aggregate, bounded rationality and the ability to learn from suboptimal decisions in the past. Interestingly, the buyers of the more expensive BC50 made fewer mistakes than those of the cheaper BC25, whose buyers lost on average. Perhaps the BC25 is cheap enough that it is not worthwhile to bother about it.

Analyzing real transactional data, our study provides valuable insights into customers' contract choice and optimization behavior. Going into more detail, further research should scrutinize the cost side of customer decision making. As long as we do not know more about consumers' cognitive costs, it will remain difficult to determine the adequate point of comparison, as our analysis has shown. This might be illustrated by the example of considering upper boundaries, i.e. over-usage becomes a mistake, too: As BahnCards can be upgraded virtually anytime throughout contract duration, customers are forced to optimize *continuously* during the year; when only considering lower boundaries customers must make only *two* decisions, one at the beginning and another one at the end of the period. Hence, taking upper boundaries into account implies that the number of time consuming and rationally demanding (i.e. costly) decisions increases dramatically. Moreover, as decisions get more complicated, the point where the marginal costs of additional information gathering equate its (expected) marginal benefits could be reached very soon, making results that are inefficient in a frictionless world the optimal ones in a second best world.

While we concentrated on the customers of BahnCards so far, there are some interesting implications for the selling side, in this case the DB AG, too. Offering BahnCards seems to be a very good idea for this service firm because the flat-rate bias brings a double dividend. Both the BahnCards and the reduced pay-per-mile fees bring in money and the net effect is clearly positive. Lower fees per mile encourage more traffic and the BahnCard fees more than compensate

the lower price per mile whereas the number of miles can be expected to be higher together with the total revenues. We do not know the costs of DB AG but they are probably mainly fixed such that the profits are up, too.

Another interesting question for future research is the optimal structure of BahnCards, both for the customers and the seller. Would more BahnCards, e.g. a BahnCard75 or a BahnCard10 be a good idea? What about the current idea of DB AG to offer an additional BC25 for a shorter time period (four months instead of one year)? Rational customers could only win and never lose by more alternatives, because users can better adapt contract choice to individual travel habits. However, the outcome is less clear for bounded-rational customers. For instance, more alternatives come with a narrowing of optimality intervals. This puts more weight on the appropriate assessment of future consumption and implies a higher probability of error. Sellers can win by more price discrimination but there are also potential downsides. For example, customers might take the anticipated regret of an erroneous decision into account, which – in the worst case – might dissuade them from choosing any option with increasing likelihood of failure. Furthermore, aspects such as e.g. raising costs of segmentation need to be considered, in order to obtain a comprehensive picture of the outcome of more price discrimination.

7. References

Arkes, H.R. and Blumer, C. (1985). The Psychology of Sunk Cost. *Organizational Behavior and Decision Process*, 35: 124-140.

Carbajo, J.C. (1998). The Economics of Travel Passes; Non-Uniform Pricing in Transport. *Journal of Transport Economics and Policy*, 22: 153-173.

Cohen-Cole, E. and Zanella, G. (2009). Unpacking Social Interactions. *Economic Inquiry,* 46: 19-24.

Conover, W.J. (1999). *Practical Nonparametric Statistics.* 3rd. ed. New York, NY: John Wiley & Sons.

Daniel, W.W. (2000). *Applied Nonparametric Statistics.* 2nd. ed. Boston, MA: PWS-Kent.

DellaVigna, S. and Malmendier, U. (2004). Contract Design and Self-Control: Theory and Evidence. *Quarterly Journal of Economics,* 119: 353-402.

DellaVigna, S. and Malmendier, U. (2006). Paying Not to Go to the Gym. *American Economic Review*, 96: 694-719.

FitzRoy, F. and Smith, I. (1999). Season Tickets and the Demand for Public Transport. *Kyklos,* 52: 219-238.

Fleiss, J.L., Levin, B. and Paik, M.C. (2003). *Statistical Methods for Rates and Proportions.* 3rd. ed. New York, NY: John Wiley & Sons.

Gibbons, J.D. and Chakraborti, S. (2003). *Nonparametric Statistical Inference.* 4th. ed. New York, NY: Marcel Dekker.

Gourville, J.T. and Soman, D. (1998). Payment Depreciation: The Behavioral Effects of Temporally Separating Payments From Consumption. *Journal of Consumer Research,* 25: 160-174.

Greene, W.H. (2008). *Econometric Analysis.* 6th. ed. Upper Saddle River (NJ): Prentice Hall.

Hair, J.F., Black, W.C., Babin, B.J. and Anderson, R.E. (2009). *Multivariate Data Analysis.* 7th. ed. Upper Saddle River, NJ: Prentice Hall.

Kling, J.P. and Ploeg, S.S. vd (1990). Estimating Local Elasticities with a Model of Stochastic Class of Service and Usage Choice. In Fontenay, A.de, Shugard, M.H. and Sibley, D.S.

(Eds.) *Telecommunications Demand Modeling. An Integrated View*. Amsterdam: North Holland.

Kridel, D.J., Lehman, D.E. and Weisman, D.L. (1993). Option Value, Telecommunication Demand, and Policy. *Information Economics and Policy,* 5: 125-144.

Lambrecht, A. and Skiera, B. (2006). Paying Too Much and Being Happy About It: Existence, Causes and Consequences of Tariff-Choice Biases. *Journal of Marketing Research,* 43: 212-223.

Lumley, T., Diehr, P., Emerson, S. and Chen, L. (2002). The Importance of the Normality Assumption in Large Public Health Data Sets. *Annual Review of Public Health,* 23: 151-169.

McAfee, R.P., Mialon, H.M. and Mialon, S.H. (2010). Do Sunk Costs Matter?" *Economic Inquiry*, 48: 323-336.

Miravete, E.J. (2002). Estimating Demand for Local Telephone Service with Asymmetric Information and Optional Calling Plans. *Review of Economic Studies,* 69: 943-971.

Miravete, E.J. (2003) Choosing the Wrong Calling Plan? Ignorance and Learning. *American Economic Review*, 93: 297-310.

Mitchell, B.M. and Vogelsang, I. (1991). *Telecommunication Pricing: Theory and Practice.* Cambridge, UK: Cambridge University Press.

Nunes, J.C. (2000). A Cognitive Model of People's Usage Estimations. *Journal of Marketing Research,* 37: 397-409.

Prelec, D. and Loewenstein, G. (1998). The Red and the Black: Mental Accouting of Savings and Debt. *Marketing Science*, 17: 4-28.

Siegel, S. and Castellan, N.J. (1988). *Nonparametric Statistics for the Behavioral Sciences.* New York, NY: McGraw-Hill.

Soman, D. (2001). Effects of Payment Mechanism on Spending Behavior. The Role of Rehearsal and Immediacy of Payments. *Journal of Consumer Research*, 27: 460-474.

Soman, D. and Gourville, J.T. (2001). Transaction Decoupling: How Price Bundling Affects the Decision to Consume. *Journal of Marketing Research*, 38: 30-44.

Steimez, S.S.C. (2009). White-Knuckle Externalities. *Economic Inquiry*, 47: 304-316.

Tamhane, A.C. (1977). Multiple Comparisons in Model I One-Way ANOVA with Unequal Variances. *Communications in Statistics,* 6: 15-32.

Thaler, R.H. (1980). Towards a Positive Theory of Consumer Choice. *Journal of Economic Behavior and Organization,* 1: 39-60.

Thaler, R.H. (1985). Mental Accounting and Consumer Choice. *Marketing Science,* 4: 199-214.

Thaler, R.H. (1999). Mental Accounting Matters. *Journal of Behavioral Decision Making,* 12: 183-206.

Train, K.E. (1991). *Optimal Regulation. The Economic Theory of Natural Monopoly.* Cambridge, MA: MIT Press.

Train, K.E., Ben-Akiva, M. and Atherton, T. (1989). Consumption Patterns and Self-Selecting Tariffs. *Review of Economic and Statistics,* 71: 62-73.

Train, K.E., McFadden, D.L. and Ben-Akiva, M. (1987). The Demand for Local Telephone Service: A Fully Discrete Model of Residential Calling Patterns and Service Choices. *Rand Journal of Economics*, 18: 109-123.

Van Vuuren, D. and Rietveld, P. (2002). The Off-Peak Demand for Train Kilometers and Train Tickets. *Journal of Transport Economics and Policy,* 36: 49-72.

Wertenbroch, K. (1998). Consumption Self-Control via Purchase Quantity Rationing. *Marketing Science*, 17: 317-337.

Winer, R.S. (2005). *Pricing.* Cambridge, MA: Marketing Science Institute.

V. ERKLÄRUNG

Ich versichere an Eides statt, dass ich die eingereichte Dissertation “Structural Market Changes and Strategic Adaptation Along the Value Chain: Theoretical Perspectives and Empirical Evidence” selbstständig verfasst habe. Andere als die von mir angegebenen Quellen und Hilfsmittel habe ich nicht verwendet. Alle wörtlich oder sinngemäß den Schriften anderer Autoren entnommenen Stellen habe ich durch Angabe der entsprechenden Quellen kenntlich gemacht. Ich versichere auch, dass diese Dissertation nicht bereits anderweitig als Prüfungsarbeit vorgelegen hat.

Carl Hendrik Schmale

Münster, 05.05.2010