Essays in Applied Microeconomics

Inauguraldissertation
zur Erlangung des akademischen Grades
eines Doktors der Wirtschaftswissenschaften
der Universität Mannheim

vorgelegt von

Isabel Ruhmer

Mannheim, März 2012

Abteilungssprecher:	Prof. Dr. Martin Peitz
Referent:	Prof. Dr. Martin Peitz
Korreferentin:	Prof. Dr. Irene Bertschek
Datum der mündlichen Prüfung:	27. März 2012

To my parents

Acknowledgements

More people have supported me in the course of the past years than could be included in this whole volume, therefore I can only thank a selection of them here.

First of all, I would like to thank my supervisor Martin Peitz for his great support. He always encouraged me to come up with my own ideas and at the same time provided invaluable guidance and feedback whenever I felt lost.

I am also grateful to Irene Bertschek for her insightful comments and for her immediate willingness to join my thesis committee.

I have always felt very privileged to be part of the CDSE program and to work in the thriving academic environment of Mannheim University. I gratefully acknowledge the financial support from the Deutsche Forschungsgesellschaft that allowed me to focus on my research in all these years. Moreover, I wish to thank Patrick Rey for inviting me to Toulouse School of Economics where I had the chance to work on the first chapter of this thesis and discuss my thoughts with experts in the field of two-sided markets, in particular with Bruno Jullien.

The biggest thanks go to my two co-authors Pia Dovern-Pinger and Johannes Koenen - working with you has been so much fun! Pia showed me how an empirical economist tackles problems and always came up with great out-of-the-box thinking. Johannes' enthusiasm for doing research and his never-ending willingness to help me answering difficult questions in all of my research projects is just incredible.

In addition, I wish to thank my fellow graduate students at the CDSE – in particular the class of 2007 – for the great time we had together. Moreover, I want to say thank you to the crowd "living" on the 3rd floor of the econ department for the friendly and inspiring atmosphere that they create. In particular, I had an awful lot of fun with Petra Loerke, Christian Lambertz and Christian Michel - thanks for being such cheerful mates.

Finally, yet foremost, my warmest thanks go to my family for their everlasting encouragement and love. My mother, who I could always call in the midst of a crisis knowing that she will make me feel better. My father, who gave me a jump-start whenever I made a big new step in my life. At last, I would like to thank Julian for never stopping to believe that I can do this and for always being there for me with love, patience and a great sense of humor.

Contents

Chapter 1

General Introduction

This dissertation consists of three self–contained papers, which contribute to different strands of the literature on industrial organization and microeconomic theory. In Chapter 2, I analyze the sustainability of price collusion in two-sided markets that are characterized by indirect network externalities between the two distinct customer groups. Chapter 3, which is joint work with Johannes Koenen, examines the incentives to innovate in a 3-player vertical industry structure with overlapping product generations. Chapter 4 empirically investigates the existence of compromise effects when people choose their main course from a restaurant's menu. It has been written jointly with Pia Dovern-Pinger. The appendix of this dissertation contains the appendices of the respective papers including mathematical proofs, model extensions, data description and regression tables. References of all chapters are found in the bibliography at the end of this thesis.

Chapter 2:
Platform Collusion in Two-Sided Markets

This chapter analyzes price collusion between platforms in a two-sided market model. Given that media markets such as newspapers or TV channels have been particularly prone to collusion, I use Armstrong's (2006) well-established model of platform competition as the stage game of an infinitely repeated price game where defection from coordinated behavior will be punished using grim-trigger strategies. I show that higher indirect network externalities have two opposing effects on the sustainability of a cartel. First, if one side values members on the other market side more highly, Nash prices fall because competition for this side gets harsher. As a consequence, punishment profits are a falling function of increasing network effects. In addition, consumers' utility from platform participation increases if they enjoy a larger benefit from the presence of platform members on the opposing market side. Hence, two-sided platforms make larger gains from collusion as network externalities grow. Second and countervailing, however, platforms also earn larger profits from deviation as network effects become stronger - a result which is due to the fact that indirect network externalities induce a feedback loop leading to more sensitive demand reactions in response to a price decrease. Comparing these opposing effects, I show that the latter effect always dominates in Armstrong's (2006) model. The reason being that a platform benefits more strongly from the existence of higher network effects under defection because of the resulting asymmetric platform sizes.

In addition to this main result, I show that firms also have an incentive to collude only on one side of the market, contradicting the hypothesis of Evans and Schmalensee (2008). In particular, if platforms fix prices on the market side where consumers enjoy higher network effects, they compete away some, but not all, of the supra-competitive profits by setting prices below the competitive level on the other side. If, however, platforms are able to form a cartel on the side of the market that suffers from negative externalities, they might actually make larger additional profits than under cartelization of both market sides.

Chapter 3:
Complementary inputs and the incentives for upstream firms to patent and innovate

This chapter studies the impact of patent protection on upstream innovation incentives in a vertical industry with complementary inputs and consecutive investment periods. In more detail, we investigate a contract theoretic setting in the spirit of

Grossman and Hart (1986) and Hart and Moore (1990), in which two upstream firms make investments into components that are perfect complements to produce a final product in each of the two periods. As a new twist, the final product to be sold in the second period can be composed of one of each components from either period - that is, a new component of one type can be substituted by the old component of the same type. We show that in case of non-cooperative, fixed-order sequential bargaining between the two suppliers and the downstream firm, ironclad intellectual property rights that protect first period investments lead to a complete breakdown of investments into components due to hold-up problems, despite full bargaining power of the investing parties. Knowledge diffusion that allows the downstream firm D to buy an older version of the component from a competitive fringe at period two is thus beneficial for all firms. In particular, knowledge diffusion generates an endogenous outside option for D in the second period, allowing each of the upstream firms to claim a part of the total surplus that corresponds to the value of their contributions to the respective coalition with D. The ability of the residual claimant to exploit the holdup-situation is thereby reduced. As a robustness check, we show that allowing for a stochastic bargaining sequence alleviates the complete hold-up under patent protection. Yet, upstream innovation might still be higher under knowledge diffusion.

In addition to these main findings, we show how our setting with two complementary investments in two periods relates to more "standard" investment setups by studying the effects of both vertical and horizontal integration. Partial vertical integration of D with one of its suppliers does not solve the hold-up with ironclad patents, but leads to higher investment levels than non-integration in case of technology diffusion. Finally, the only situation leading to first-best investment levels is when both upstream suppliers merge and first-period innovations are protected by long-lasting IP rights, a solution which resembles the classical result of Grossman and Hart (1986).

Chapter 4:

The Compromise Effect in Action: Lessons from a Restaurant's Menu

Standard discrete choice analysis assumes that consumers behave rationally when choosing an option from a given set, implying that a change in choice set composition does not lead to choice reversal between options available in both sets. Many laboratory experiments, however, have shown that individuals do not necessarily behave according to standard theory when facing changes in the choice set. In particular, one of the most well-known context effects is the so-called "compromise effect", documented first by Simonson (1989). It refers to a situation where a choice alternative gains market share when it moves from being an extreme option to a compromising or middle option with respect to all relevant choice attributes. This article contributes to the existing literature by being the first to provide "real world" evidence for the existence of a compromise effect. To this end, we construct a completely new and unique data set from raw data provided by a German specialties restaurant. Overall, we observe more than 88,000 individual choices of main courses from 21 different menus offered over a period of more than 7 years. Variation in price, position and number of items presented within 6 different main dish categories from one menu to another allows us to investigate the effects of several "pure" cases of choice set expansion as well as to estimate various discrete choice models. Our findings indicate that the compromise effect prevails both in descriptive and regression analyses. It is more pronounced when an alternative switches from being the most expensive item (instead of being the lowest-price item) in the choice set to being a compromise option. Moreover, our results show that controlling for confounding factors such as the background context of individual decision makers is important, as it can change the size of the compromise effect.

Chapter 2

Platform Collusion in Two-Sided Markets

2.1 Introduction

Price-fixing cartels have been observed in various two-sided industries.[1] For instance, the two major arts auction houses Christie's and Sotheby's fixed both seller's commission fees and trading conditions for buyers for almost seven years until their cartel was uncovered by competition authorities.[2] The practice of payment card associations to fix interchange fees has steered the focus of various antitrust authorities. In the Netherlands, for example, the joint venture Interpay and its member banks were fined for "charging excessive rates" in 2004. Moreover, both the Reserve Bank of Australia (in 2002) and the Banco de Mexico (in 2004) forced payment card issuers to reduce their interchange fees after thorough investigation of the industry (for details see e.g. Negrín (2005), Weiner and Wright (2005)). Even the academic

[1]See Evans and Schmalensee (2008) for a summary of antitrust cases in two-sided markets.

[2]The European Commission stated that agreements were foremost related to the conditions applicable to sellers, but they also showed that conditions for buyers were included in the cartel rules. Hence, both sides of the market were involved in the collusive practice. After a detailed investigation, the European Commission considered both auction houses guilty of a serious infringement of Article 81(1) of the EC Treaty, which by its very nature led to an important distortion of competition to the exclusive benefit of the participating companies and to the detriment of customers. In effect, an initial fine of over EUR 40 million was imposed, but - due to leniency program application - reduced to zero fines for Christie's and around EUR 24 million for Sotheby's. For further details on the case, see European Commission (2002).

market is prone to collusion. In fact, 23 U.S. universities colluded on financial aid awards for students for over 40 years. Their practices eventually ended after the Department of Justice charged the participating Ivy League universities and the MIT of illegal price fixing in 1991 and an out-of-court settlement forbidding discussion of and coordination on prospective awards was signed by all members. An interesting detail which makes this case a suspect for cartelization on both sides of the academic job market is that press reports and evidence presented in court indicated that collusion on faculty salaries may have occurred as well.[3] Finally, joint price-fixing has been observed in various media markets. For instance, the two dominating players in the German private TV market, ProSiebenSat1 and the RTL-Group, have repeatedly been under investigation for coordinated behavior in the advertising market. In 1999, after a striking convergence of advertisement prices and the simultaneous announcement of a price increase which was justified by almost identical wording, the German competition authority opened a first investigation (see Budzinski and Wacker (2007) for details). In 2007, ProSiebenSat1 and the RTL-Group had to pay high fines and change the pricing of their commercial airtime after an investigation by the Bundeskartellamt (2009). Another case involved three German nationwide newspaper publishers that planned to build a common agency for employment advertisements.[4] Their contract included fixed prices for employment ad space and explicit profit sharing rules.[5] Moreover, newspaper cartels that were fixing cover prices have been reported in Australia and Switzerland (Merrilees (1983), Wettbewerbskommission der Schweizerischen Eidgenossenschaft (2000a,b)). Finally, newspapers both in the Netherlands and the U.S. were legally allowed to fix both cover and ad space prices jointly under the Dutch Newspapers Publishers Association and the Newspaper Preservation Act, respectively (see Van Kranenburg, Palm, and Pfann (2002); Van Kranenburg (2001); Romeo, Pittman, and Familant

[3]For more details see e.g. Masten (1995) or Salop and White (1991). I would like to thank Patrick Rey for mentioning this example of a two-sided cartel to me.

[4]Those publishers were Süddeutsche Zeitung GmbH, Druck- und Verlagshaus Frankfurt am Main GmbH and Axel Springer Verlag AG.

[5]The German competition authority banned the common agency. After Axel Springer Verlag AG dropped out of the agreement, the final court of appeal allowed the remaining parties to form the joint agency. For more details on the case, consider the report of the German competition authority (Bundeskartellamt, 1999) or an article in the "Handelsblatt" from July 22, 2002 (Handelsblatt, 2002) which summarizes the proceedings of the investigation and the final court decision.

(2003); Picard (1995)).

As these cases indicate, there is scope for different forms of collusive behavior in two-sided markets. Especially in case of media markets, one can observe both collusion on just one side of the market, either readers or advertisers, as well as "full" collusion on both sides. Moreover, it might be that platforms cannot choose on which side to collude, e.g. in case of free-to-air TV channels which earn all profits by selling air time to advertisers. The unifying element of all these industries, however, is that their pricing strategies are influenced by the presence of two distinct consumer groups interrelated via indirect network externalities. In order to understand how and when price collusion is possible in two-sided markets, this paper analyzes the effects of these indirect network effects on collusive stability. Given that media markets seem to be very prone to price collusion, I will take a first step in understanding the incentive to collude in presence of indirect network externalities using Armstrong's (2006) model of two-sided competition, which has become the workhorse model for investigating pricing and antitrust issues in media markets.[6] Furthermore, the simple, yet very general structure of this model allows me to analyze different setups such as single- versus multi-homing consumers, paying versus non-paying readers or viewers, and collusion on just one versus both sides of the market. Taking Armstrong's (2006) model as a stage game and assuming grim trigger punishments in case of defection from the collusive agreement, I find that increasing network externalities have two opposing effects on collusive stability in all setups.[7] First, stronger network effects raise platforms' incentive to collude, namely by increasing the difference between collusive and Nash profits. If one market side values members on the opposite side more highly, Nash prices fall because competition for this side gets harsher. At the same time, consumers' utility from platform participation increases if they benefit more from the presence of platform users on the opposing market side. Consequently, platforms can earn larger collusive profits by setting monopolistic cartel prices when utility grows with rising network externalities. Second and countervailing, however, platforms

[6] See, for example, Kaiser and Wright (2006); Peitz and Valletti (2008); Choi (2010) for an application of Armstrong's framework to media markets.

[7] Please note that this means that increasingly negative externalities have the opposing effect. In other words, whenever I speak of "stronger" network effects, I *do not* refer to network effects in absolute value.

also gain larger profits from deviation when network effects become stronger - a result which is due to more sensitive demand reactions when both market sides are interlinked by indirect externalities. In fact, deviation becomes more attractive because rising network effects allow the defector to steal more and more consumers from its competitor on both sides when lowering his prices. Comparing those two opposing effects and solving for the critical discount factor, I show that the latter effect always dominates in my model. In other words, collusion becomes harder to sustain as network externalities between the market sides grow. Furthermore, I find that an increasing asymmetry in the network benefits between both sides of the market has a negative impact on collusive sustainability in case that both sides single-home, whereas an increasing asymmetry makes it easier to collude if one side multi-homes.

My results confirm Evans & Schmalensee's (2008) hypothesis that collusion is generally harder to sustain in two-sided markets. Their argument, however, is that successful cartels need to coordinate prices on both sides of the market which asks for more agreements and monitoring and makes it more difficult to form an effective cartel. My paper shows that two-sided collusion becomes harder to sustain even without increased monitoring or coordination costs. A second point made by Evans & Schmalensee concerns collusion only on one of the two market sides. They argue that such one-sided agreements do not constitute an alternative collusive scheme since all supra-competitive profits earned on the colluding side would be competed away on the other one because of feedback effects in demand.[8] Given my framework, however, I cannot bear out their statement. Although some of the collusive profits might be competed away on the remaining competitive market side if network effects are symmetric on both sides or if platforms are forced to collude on the market side that imposes lower external benefits, platforms still benefit from such a one-sided cartel.[9] Even more surprisingly, collusive profits might even be higher than under two-sided collusion if platforms are able to form a cartel on the market side that imposes higher indirect externalities. Finally, depending on which side platforms decide to collude on, the critical discount factor might be higher or

[8]Note that this statement is also falsified by the existence of one-sided collusion in real world examples, see above.

[9]This result is also confirmed by Dewenter, Haucap, and Wenzel (2011) for a different two-sided market model.

lower than under two-sided collusion.

This paper adds to the growing body of literature on two-sided markets that originates from seminal contributions by Rochet and Tirole (2003, 2006), Caillaud and Jullien (2003) and Armstrong (2006). In particular, it enriches a strand which focuses on the impact of indirect network externalities on well established competition policy results,[10] providing the first theoretical investigation of collusive sustainability in two-sided markets. In parallel work, however, Dewenter, Haucap, and Wenzel (2011) analyze welfare implications of collusion in media markets. They show that collusion only on advertisement slots might actually increase total welfare because it leads to less advertising and lower copy prices. Welfare implications of collusion on both market sides (readers and advertisers) are ambiguous, although it might still be possible that welfare increases. In contrast to the model used in this paper, they completely ignore the question of cartel sustainability. Moreover, their model makes the simplifying assumption that firms first choose advertising quantities and then compete in newspaper copy prices instead of choosing prices simultaneously for both sides. Finally, there exists one empirical paper by Argentesi and Filistrucchi (2007) that tests for the presence of collusion in media markets using a structural model. They analyze the Italian newspaper market and address the question whether observed price patterns are consistent with profit-maximizing behavior by competing firms or instead driven by some form of (tacit or explicit) coordinated practice. Their model encompasses a demand estimation for differentiated products on both sides of the market and allows for profit maximization by the publishing firms taking into account the possible interconnections between readers and advertisers. In order to simplify the analysis, they assume that readers do not care about the amount of advertisement to be found in newspapers, i.e. indirect network effects are present only in one direction. They derive hypothetical markups under the two alternative conjectures of competition and joint profit maximization and compare them with actually observed ones. Using this method, they find an

[10] Areas of competition policy that have been covered by recent papers include mergers, tying and bundling, exclusive contracts and price discrimination. See, for example, Chandra and Collard-Wexler (2009), Armstrong and Wright (2007), Rochet and Tirole (2008) or - for an overview - Evans (2003), Evans and Schmalensee (2008) as well as Rysman (2009). Moreover, White and Weyl (2010) propose a general framework for imperfect platform competition allowing, e.g., first-order merger analysis.

indication of joint profit maximization for cover prices, whereas the advertising market is closer to competition, a result which they claim to be consistent with anecdotal evidence of the Italian newspaper market.

The rest of the paper is organized as follows: In the next section, I will outline the framework and describe the collusive game. Section 2.3 presents and discusses the effects of increasing network externalities on collusive stability if both sides of the market single-home, i.e. buy from just one platform. To this end, I first consider the case where platforms have the possibility to fix prices on both market sides (see subsection 2.3.1), before turning to the analysis of so-called one-sided collusion where platforms coordinate only on one side of the market (subsection 2.3.2). Section 2.4 is devoted to the analysis of collusion when consumers have the possibility to join multiple platforms at the same time, i.e. when they are able to multi-home. Finally, section 2.5 concludes.

2.2 The Model

The stage game of my infinitely repeated collusive setup is based on the refined version of Armstrong's (2006) famous two-sided market model, as presented in Armstrong and Wright (2007). The framework fits well to media markets in general and - to some extent - even to the academic market example given in the introduction.[11] In particular, pricing for the service offered by a platform is based on membership fees rather than on a transaction-based payment. In case of media markets, this corresponds to the fact that readers and advertisers "meet" in a newspaper or magazine, but the price that they pay for getting access to the other side does not (directly) depend on how many agents they will actually reach. Furthermore, the imposed network externalities interconnecting both market sides are assumed to be purely membership-based, i.e. each consumer on one market

[11] Although Armstrong's model fits the Ivy League cartel in most aspects, I have to admit that the indirect network benefits that students or faculty members enjoy at a certain university are not purely membership-based. Instead, the academic excellence of university members on either side of the academic market highly matters. Such a quality dimension, however, is not included in Armstrong's framework.

side values the per-se presence of members on the other side.[12]

There are two platforms, e.g. newspapers or TV channels, denoted by A and B, which are located at the endpoints of a Hotelling line with unit length. These platforms serve two types of customers, group 1 and group 2. Each group has a unit mass of agents who are uniformly distributed along the unit line. Each agent is interested in purchasing a single unit from a platform. If a customer of group i located at $x \in [0,1]$ decides to join platform A or B, she receives the following respective utility:

$$u_i^A = k + a_i n_j^A - p_i^A - tx \ ; \ u_i^B = k + a_i n_j^B - p_i^B - t(1-x) \quad \text{with } i \neq j;\ i,j \in \{1,2\} \tag{2.1}$$

where k is the intrinsic benefit from joining a platform, e.g. the utility that a reader or viewer enjoys from content provided by a newspaper or TV channel.[13] Further, a_i with $i \in \{1,2\}$ is the benefit that an agent of group i enjoys from the presence of each agent on the other market side,[14] while p_i^j describes the lump-sum price that an agent of group i pays to platform $j \in \{A,B\}$ when she joins and t is the transportation cost parameter measuring horizontal product differentiation between the two platforms.[15]

In the following sections, I will investigate the effects of collusion in two different demand setups. First, I will consider the case where agents from both groups single-home, i.e. where they can either buy from A or B. In line with Armstrong (2006), I assume k to be high enough to guarantee that all agents on either market side wish to subscribe to a platform in equilibrium, which implies that $n_1^A = 1 - n_1^B$

[12]There exist other forms of network externalities which might be present in two-sided markets (see e.g. Weyl (2010)). Moreover, a platform might apply alternative payment schemes. Thus, my analysis is only a first step in understanding the sustainability of collusive practices in two-sided markets.

[13]I assume that the intrinsic utility of joining a platform is identical on both market sides. My main results, however, are robust to a generalization with different intrinsic utility levels on each side. A detailed proof is available upon request.

[14]For a possible micro foundation of this reduced-form network benefit structure, see Belleflamme and Peitz (2010).

[15]The results presented in section 2.3.1 are robust against both a switch to quadratic transport costs as well as to allowing for different transport costs on each market side.

and $n_2^A = 1 - n_2^B$.[16] As a result, the number of agents from each group that join platform i under single-homing follows from solving for the indifferent consumer on each side:

$$n_1^i = \frac{1}{2} + \frac{a_1(2n_2^i - 1) + (p_1^j - p_1^i)}{2t} \; ; \; n_2^i = \frac{1}{2} + \frac{a_2(2n_1^i - 1) + (p_2^j - p_2^i)}{2t} \qquad i, j \in \{A, B\} \; ; \; i \neq j \tag{2.2}$$

Given the price pairs (p_1^A, p_2^A) and (p_1^B, p_2^B) offered by the platforms A and B, respectively, these simultaneous demand functions solve for the following market shares:

$$n_1^i = \frac{1}{2} + \frac{a_1(p_2^j - p_2^i) + t(p_1^j - p_1^i)}{2(t^2 - a_1 a_2)} \; ; \; n_2^i = \frac{1}{2} + \frac{a_2(p_1^j - p_1^i) + t(p_2^j - p_2^i)}{2(t^2 - a_1 a_2)} \qquad i, j \in \{A, B\} \; ; \; i \neq j \tag{2.3}$$

Second, I turn to the case where one market side is allowed to buy from both platforms while the other side still single-homes.[17] In order to make this setting both comparable to the single-homing setup and as tractable as possible, I will assume that a multi-homing agent's utility simply corresponds to $u_j^A + u_j^B$. In consequence, her decision boils down to buying from platform i if $u_j^i \geq 0$.[18] Platform A faces the following demand on the multi-homing side:

$$n_i^A = \frac{k + a_i n_j^A - p_i^A}{t} \quad \text{with } i, j = \{1, 2\}, i \neq j \tag{2.4}$$

[16] The exact assumption on k will be stated later on.

[17] As pointed out by Armstrong (2006); Armstrong and Wright (2007), if agents on one side multi-home, there is no incentive for agents on the other market side to do so as well.

[18] While I simply assume that agents can either single- or multi-home, Armstrong and Wright (2007) show that it depends upon the size of network benefits relative to transport costs if consumers actually wish to multi-home. A restriction of their multi-homing framework, however, is that it is not directly comparable to the single-homing setup with respect to transport costs and intrinsic utility. In particular, they assume that $t = 0$ on the multi-homing side and that consumers receive k only once even if joining both platforms. In consequence, there are only four possible demand configurations on the multi-homing side: all agents multi-home, all single-home either on platform A or B or no agent joins any platform. Moreover, their setup is problematic in so far as there are ranges of prices where more than one of these configurations are consistent. Hence, multiple Nash equilibria might arise.

while demand for platform B is given by:

$$n_i^B = \frac{k + a_i n_j^B - p_i^B}{t} \quad \text{with } i,j = \{1,2\}, i \neq j \tag{2.5}$$

The number of multi-homing agents on side i will therefore be equal to the maximum of $\{(n_i^A + n_i^B - 1), 0\}$.

In line with Armstrong (2006), I will make the assumption that network externality parameters a_1 and a_2 are small enough in comparison to the differentiation parameter t such that a market-sharing equilibria always exists. In other words, the following sufficient and necessary condition must be fulfilled to guarantee that both platforms will be active in the competitive equilibrium instead of having one platform attracting all agents:

Assumption 2.1 $4t^2 > (a_1 + a_2)^2$

Finally, I will assume that platforms' costs of production on either side are normalized to zero to focus on collusive sustainability as a function of indirect network effects only.

Given the above described stage game, I analyze a standard infinitely repeated price game where platforms choose their prices on both sides simultaneously in each period and discount their profits with a common discount factor δ. To evaluate the sustainability of collusion, I derive the critical discount factor above which a collusive agreement on monopoly prices can be supported by a grim trigger strategy.[19] In case of two-sided markets, such a grim-trigger strategy implies that platforms set monopoly prices on one ("one-sided collusion") or even both market sides ("two-sided collusion") and stick to them as long as no deviation is observed. In case of defection, Nash reversion on both market sides is assumed. Thus, the critical discount factor $\hat{\delta}$ equates the profits earned from sticking to the collusive agreement every period and the profit stream in case of defection, namely the sum of the optimal defection profit in the period of deviation and the stream of Nash profits in all

[19] As introduced by Friedman (1971), this strategy states that firms set prices at the monopoly level in period 1 and stick to these prices in all following periods as long as both firms adhered to the agreement in the past. If either firm deviates from this collusive price, both immediately revert to the static Nash equilibrium in prices and stick to it forever.

periods afterwards. Formally speaking, $\hat{\delta}$ solves the following incentive constraint of a colluding platform i:[20]

$$\underbrace{\frac{1}{1-\delta}\pi_i^C}_{\text{present value of collusive profits}} \geq \underbrace{\pi_i^D + \frac{\delta}{(1-\delta)}\pi_i^N}_{\text{present value of optimal defection}}$$

$$\Leftrightarrow \quad \delta \geq \hat{\delta} \equiv \frac{(\pi_i^D - \pi_i^C)}{(\pi_i^D - \pi_i^N)} \qquad i \in \{A, B\} \tag{2.6}$$

For all discount factors bigger than $\hat{\delta}$, platforms will find it more profitable to collude on monopoly prices than to deviate. Hence, if $\hat{\delta}$ increases in response to an increase of indirect network externalities, then collusion becomes harder to sustain as consumers on one side value consumers on the other side more strongly. It is important to note that the overall effect of a_i on $\hat{\delta}$ needs to be disentangled to fully understand its mechanism. Therefore, I will separately look at the effect of a_i on a platform's gain from colluding and on its incentive to deviate.

2.3 Single-Homing

Let me start my analysis by investigating different forms of coordinated behavior in a setup where both consumer groups single-home. The empirical relevance of single-homing has been investigated by Kaiser and Wright (2006). Testing different versions of Armstrong's model on data for the German magazine market, they concluded that competition between platforms is prevalent on both market sides. Put differently, the assumption of multi-homing on the advertisers' side does not provide a good fit of the German magazine market. Instead, both readers and advertisers opt for just one magazine. In a broader sense, real-world examples of single-homing environments might also be motivated by indivisibilities, limited resources or contractual restrictions. Thus, this setup matches both the Ivy League case as well as the cartel of German nationwide newspaper publishers mentioned in the introduction.

[20] Note that for a one-sided market, i.e. two horizontally differentiated firms selling only one product to one group of customers, this collusive game has been analyzed by Chang (1991) and Häckner (1995).

In the following, I will first investigate collusive sustainability in case that platforms coordinate their pricing on both market sides before turning to the situation where collusion is possible only on one of the two sides.

2.3.1 Two-Sided Collusion

Given the infinitely repeated game described in section 2.2, each platform's punishment profit is equal to its stage game Nash profit. Using market shares as given in (2.3), platform i's profit is equal to $\pi_i = p_1^i n_1^i + p_2^i n_2^i$. Maximizing profits with respect to p_1^i and p_2^i and rearranging first order conditions, I obtain:

$$p_1^N = t - \frac{a_2}{t}\left(a_1 + p_2^N\right) \;\; ; \quad p_2^N = t - \frac{a_1}{t}\left(a_2 + p_1^N\right) \tag{2.7}$$

Note that the classical Hotelling price on each side, which would be equal to t, is reduced by the external benefit that a platform enjoys from attracting one additional consumer on this side. Solving those two simultaneous equations yields the symmetric Nash equilibrium prices $p_1^N = t - a_2$ and $p_2^N = t - a_1$. Thus, one side of the market will be targeted more aggressively than the other if that side's consumers impose larger external benefits on the other side's consumers than vice versa.

In order to guarantee that both market sides are fully covered given the Nash equilibrium prices of the stage game, I will make the following assumption throughout the remaining analysis:[21]

Assumption 2.2 $k \geq \max\{\frac{3}{2}t - \frac{a_1}{2} - a_2, \frac{3}{2}t - \frac{a_2}{2} - a_1\}$

Under this assumption, each platform gains a fifty percent market share on both sides and Nash profits are given by:

$$\pi^N = \frac{1}{2}p_1^N + \frac{1}{2}p_2^N = t - \frac{a_1 + a_2}{2} \tag{2.8}$$

[21] If assumption 2.2 is fulfilled, all consumers will find it optimal to join one of the two platforms under Nash prices, given that their utility of not joining a platform is normalized to zero.

Hence, punishment profits fall as network externalities a_1 or a_2 increase, while they increase in product differentiation t.

The maximum collusive prices platforms can set in a cartel correspond to those prices that set the indifferent consumer's utility equal to zero on both sides. Since platforms are located symmetrically, monopoly prices differ only with respect to network externality parameters, namely $p_i^C = k - \frac{t}{2} + \frac{a_i}{2}$ with $i = 1, 2$. Collusive profits follow directly as symmetrically located firms will split both market sides equally:

$$\pi^C = \frac{1}{2}p_1^C + \frac{1}{2}p_2^C = k - \frac{t}{2} + \frac{a_1 + a_2}{4} \tag{2.9}$$

Summing up, it is easy to see that a platform's gain from colluding $(\pi^C - \pi^N)$ is increasing in both network externality parameters and the following lemma can be stated.

Lemma 2.1 *A two-sided platform's incentive to collude* $(\pi^C - \pi^N)$ *becomes stronger if* a_i*, with* $i \in \{1, 2\}$*, increases. First, a higher* a_i *reduces Nash prices and profits to be earned on the opposing market side* j*. Second, a higher benefit* a_i *allows for a higher maximum collusive price and thereby higher profits earned on market side* i*.*

In the newspaper industry, this result would imply that competing newspapers earn higher collusive profits if network effects grow, because market cartelization enables them to skim away all utility from readers and advertisers. The more an advertiser cares about the number of readers seeing his advertisement, the higher his reservation price. The less readers dislike ads (or the more they like them) the higher is the cover price they are willing to pay. In contrast, competitive Nash profits are decreasing in the indirect network effects because platforms price in any external benefit that they earn from attracting additional consumers. If, for example, readers become more valuable to advertisers, newspapers will demand a lower cover price in order to attract more readers. Overall, newspapers' incentives to collude are higher the more advertisers and readers value each other.

Countervailing the incentive to collude, firms always have an incentive to deviate from the collusive agreement in order to earn larger stage-game profits. The optimal defection strategy of platform i is given by its price reaction functions R_1^i and R_2^i:

$$R_1^i = \begin{cases} \frac{p_1^j+t}{2} + a_1\left(\frac{p_2^j-p_2^i}{2t}\right) - \frac{a_2}{t}\left(\frac{a_1+p_2^i}{2}\right) & \text{if } p_1^j < \overline{p}_1^j \\ p_1^j - t + \frac{a_1 a_2}{t} + \frac{a_1}{t}(p_2^j - p_2^i) & \text{if } p_1^j \geq \overline{p}_1^j \end{cases} \tag{2.10}$$

$$R_2^i = \begin{cases} \frac{p_2^j+t}{2} + a_2\left(\frac{p_1^j-p_1^i}{2t}\right) - \frac{a_1}{t}\left(\frac{a_2+p_1^i}{2}\right) & \text{if } p_2^j < \overline{p}_2^j \\ p_2^j - t + \frac{a_1 a_2}{t} + \frac{a_2}{t}(p_1^j - p_1^i) & \text{if } p_2^j \geq \overline{p}_2^j \end{cases} \tag{2.11}$$

$$\text{with } i,j \in \{A,B\}\ ;\ i \neq j\ ;\quad \overline{p}_1^j = 3t - \frac{(a_1+a_2)^2}{2t} - \frac{(a_1+a_2)(t+a_1+p_2^j)}{2t};$$
$$\overline{p}_2^j = 3t - \frac{(a_1+a_2)^2}{2t} - \frac{(a_1+a_2)(t+a_2+p_1^j)}{2t}$$

First, note that the reaction function of platform i on side 1 depends on both prices of her competitor as well as on her own price on the opposite market side. Furthermore, both reaction functions consist of two parts as it might be an optimal reaction for platform i to monopolize one or even both market sides when prices set by the competitor are high enough. In case that monopolization is optimal, platform i will choose the lowest price that allows her to gain a market share equal to one on the corresponding market side. In the following, I interpret these reaction functions in more detail to shed some light on the influence of network effects on the optimal defection strategy. Since the reaction functions are symmetric, interpretation will be limited to the side-1 reaction function only.

Case 1 - No market monopolization If the competitor's prices p_1^j and p_2^j are low enough, platform i's optimal reaction corresponds to the first line of equation (2.10). The first part of this term, $\frac{p_1^j+t}{2}$, is equal to the classical Hotelling reaction function, i.e. the optimal price platform i imposes on side-1 customers depends positively on the differentiation parameter t and the price of her competitor on this side. In addition, however, the optimal side-1 price is influenced by the second market side via the existing indirect network externalities. *Assume for a moment* *that network externalities* a_1 *and* a_2 *are both positive.* Then, the second term of equation (2.10), $a_1\left(\frac{p_2^j-p_2^i}{2t}\right)$, indicates that an increase in the price differential between platform i and her competitor on the opposite market side increases i's price on side 1. First, a price advantage of $(p_2^j - p_2^i) > 0$ increases demand for i on side 2 by $1/2t$. This, in return, raises a consumer's utility on side 1 by a_1 times the demand increase. Hence, platform i can charge its customers on side 1 a higher price when it is less expensive than her competitor on side 2. The third term in equation (2.10) measures the external benefit of a decrease in p_1^i. Suppose that p_1^i is decreased exactly by the amount that causes an additional type-1 consumer to join i. In return, this will attract a_2/t additional type-2 consumers.[22] Those additional type-2 consumers will generate an extra revenue of p_2^i times a_2/t. Furthermore, they increase a type-1 consumer's utility by a_1 and thus the revenue to be extracted on side 1. Wrapping up, the larger the external benefit a_1 becomes, the smaller will be the optimal price reaction on side 1. *If network externalities are both negative*, then the second term of equation (2.10) will be negative if $(p_2^j - p_2^i) > 0$, i.e. a price advantage on market side 2 will decrease the optimal price on side 1. The overall influence of the third term - the external benefit - is ambiguous and depends on the size of a_1 and p_2^i. In general, the first part of equation (2.10) shows that platforms' best responses on one side might depend positively or negatively on its own price set on the opposite market side. For given rival's prices, R_1^i will be decreasing in i's side-2 price if the total external benefit $(a_1 + a_2)$ that consumers enjoy is positive. In contrast, when consumers' total network benefit is negative, i.e. when $(a_1 + a_2) < 0$, then the price reaction function for side 1 is increasing in p_2^i. Hence, a positive network externality a_1 of consumers on market side 1 implies that R_1^i is an increasing function of prices set by the competitor, but it will only increase in p_2^i if consumers on side 2 have a negative externality which is larger than a_1 in absolute terms. If consumers on side 2 also have a positive indirect network valuation, then platform i will always reduce its side-1 price in reaction to an increase in p_2^i.

Case 2 - market monopolization If the competitor's prices are so high that monopolization of market side 1 is the best response, then the second line of equa-

[22]This can be seen by differentiating n_2^i with respect to n_1^i in equation (2.2).

tion (2.10) indicates the optimal price choice. Its first part, $p_1^j - t$, is once again equal to the classical Hotelling reaction function. The second part, however, indicates that the maximum price that guarantees market monopolization on side 1 depends on the other market side as well. If a_1 and a_2 have the same sign, then p_1^i will be increased by $a_1 a_2/t$ in comparison to a market without network externalities. In addition, a positive network effect on side 1 allows to further increase the optimal price if platform i's price on side 2 is lower than the one of her competitor. As it has been the case for the optimal reaction without market monopolization, a better price on market side 2 increases demand on this side, which in return allows i to charge higher prices on side 1, too. Finally, if $a_1 > 0$, then the side-1 reaction function of platform i is decreasing in p_2^i, i.e. if side-1 consumers have a positive network benefit, then a price increase on market side 2 which causes the number of customers on this side to fall, has to be compensated by a price decrease on side 1 to avoid that the latter consumers also switch to the competitor because of their loss in utility.

From the above remarks on optimal deviation strategies, it is easy to see that the impact of network externalities on deviation prices and the resulting deviation profits is not as clear cut as its effect on collusive and Nash profits. Instead, equations (2.10) and (2.11) imply that deviation prices might be increasing or decreasing in a_1 and a_2. To figure out if deviation profits will effectively rise or fall in response to a change in network effects, the following two-step procedure is chosen: First, focus will be on the simple case when network externalities are symmetric, namely when $a_1 = a_2 = a$. Taking the results for symmetric externalities as a benchmark, the second step will be to investigate the impact of an increasing asymmetry between network effects on deviation incentives.

2.3.1.1 Symmetric Externalities

When network externalities are symmetric, i.e. when $a_1 = a_2 = a$, solving both reaction functions simultaneously under the assumption that the rival platform sticks to collusive prices p_i^C yields the following optimal deviation strategies:

$$R_i^D(p_i^C) = \begin{cases} \frac{p_i^C+t}{2} - \frac{a}{2} & \text{if } k < \overline{k} \\ p_i^C - t + a & \text{if } k \geq \overline{k} \end{cases} \quad \text{with } i = 1,2 \text{ ; } \overline{k} = \tfrac{7}{2}(t-a) \tag{2.12}$$

In essence, the standard Hotelling price reaction is decreased by the external benefit from attracting an extra agent on the opposite side. The influence of prices on the other market side as shown in equations (2.10) and (2.11), however, fully cancels out in case of symmetric externalities. Note that if collusive prices are high enough or, as stated in equation (2.12), if the intrinsic utility k is low enough compared to the difference between transport costs and network benefits, a deviating platform will find it profitable to monopolize both market sides by choosing its deviation price exactly such that an agent located at the position of the rival platform is indifferent.

Plugging those prices as well as monopoly prices for the rival platform into the demand equations given in (2.3), deviation profits are derived as follows:[23]

$$\pi_{SE}^D = \begin{cases} \frac{\left(k + \frac{t}{2} - \frac{a}{2}\right)^2}{4(t-a)} & \text{if } k < \overline{k} \\ 2k - 3t + 3a & \text{if } k \geq \overline{k} \end{cases} \tag{2.13}$$

Equation (2.13) is an increasing function of the network externality parameter a. In other words, if consumers on one side value consumers' presence on the other side more highly, platforms earn higher profits from optimal defection. The effect is rather intuitive when recalling how demand functions depend on indirect network benefits. From equation (2.2), it is easy to see that keeping prices on side 2 fixed, an extra group-1 consumer joining the deviating platform attracts a further a/t group-2 agents to the platform. These additional group-2 agents have once again a positive effect on demand of group-1 agents. Therefore, a feedback loop is started once deviation prices are slightly smaller than monopoly prices. In consequence, an optimal defection strategy leads to deviation prices that fall by

[23]Note that for symmetric network effects, Nash and collusive prices are identical on both market sides, namely $p^N = t - a$ and $p^C = k - \frac{t}{2} + \frac{a}{2}$. Hence, Nash and collusive profits are just given by their respective prices, i.e. $\pi_{SE}^N = t - a$ and $\pi_{SE}^C = k - \frac{t}{2} + \frac{a}{2}$.

less than demand for the deviating platform increases when a gets bigger. Overall, deviation thus becomes more profitable as a increases because the demand effect always dominates. Summing up the analysis of optimal defection, it is easy to show that a platform's one-time gain from defection $(\pi^D - \pi^C)$ is increasing in the indirect network externality a if assumption 2.1 is fulfilled. I can therefore state the following lemma:

Lemma 2.2 *Given that network benefits are symmetric, i.e.* $a_1 = a_2 = a$, *a platform's gain from defection* $(\pi^D - \pi^C)$ *increases in* a*. Deviation prices fall in* a*, but at the same time demand reacts with stronger positive feedback loops leading to an increase in deviation profits that outweighs the increase in collusive profits due to stronger network effects.*

Recalling Lemma 2.1, I find that increasing network benefits have two opposing effects. In consequence, the overall impact of an increase in a on collusive sustainability depends on whether the increased deviation incentive is dominated by larger gains from colluding or vice versa. Solving for the critical discount factor $\hat{\delta}_{SE}$ above which monopoly prices can be sustained as a collusive strategy in case of symmetric externalities yields:

$$\hat{\delta}_{SE} = \frac{(\pi_{SE}^D - \pi_{SE}^C)}{(\pi_{SE}^D - \pi_{SE}^N)} = \begin{cases} \frac{2k-3t+3a}{2k+5t-5a} & \text{if } k < \overline{k} \\ \frac{2k-5t+5a}{4k-8t+8a} & \text{if } k \geq \overline{k} \end{cases} \tag{2.14}$$

As comparative statics show, $\hat{\delta}_{SE}$ is increasing in a if $k > 0$.[24] I can therefore sum up the above analysis with the following proposition.

Proposition 2.1 *Suppose that assumptions 2.1 and 2.2 are fulfilled and indirect network externalities are symmetric, i.e.* $a_1 = a_2 = a$*. Then,* $\frac{\partial \hat{\delta}_{SE}}{\partial a} > 0$*. As* a *increases, the rising incentive to collude is always dominated by larger gains from optimal defection.*

[24]Since assumption 2.1 simplifies to $t^2 > a^2$ in case of symmetric externalities, it follows directly from assumption 2.2 that $k > 0$.

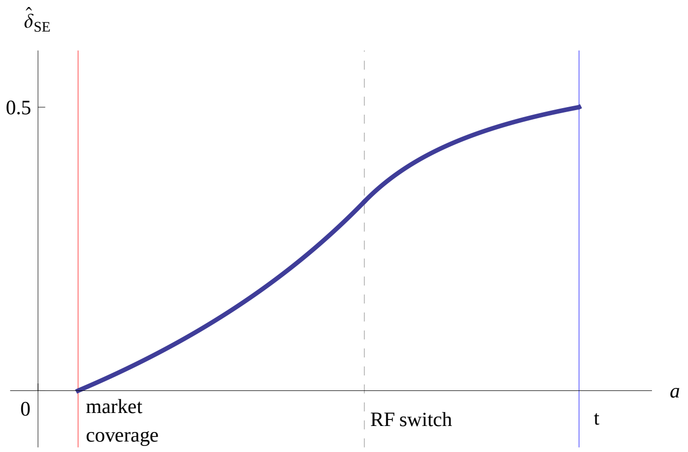


Figure 2.1: critical discount factor for symmetric externalities

For $k = 50$ and $t = 36$, the critical discount factor $\hat{\delta}_{SE}$ is plotted as a function of a in figure 2.1. The vertical line named "RF switch" refers to the value of a at which the optimal defection strategy switches from market sharing to market monopolization, namely when $k = \overline{k}$.

As this graph illustrates, $\hat{\delta}_{SE}$ increases monotonically in a. At the lowest possible value of a that still satisfies assumption 2.2 (market coverage), maximum collusive profits are equal to Nash profits and, as a natural consequence, sustainable for all discount factors between 0 and 1. Hence, $\hat{\delta}_{SE}$ is equal to zero. The maximum feasible value for a is given by assumption 2.1 which guarantees existence of a market sharing equilibrium. In case of symmetric externalities, this assumption simplifies to $t > |a|$. For $a \to t$, $\hat{\delta}_{SE}$ converges to the critical discount factor for homogeneous goods Bertrand competition, $\hat{\delta}_{SE} \to \frac{1}{2}$. Finally, the impact of symmetric network externalities on the sustainability of collusion is decreasing in the other model parameters k and t. In other words, if the intrinsic utility of platform participation grows or if both platforms become more differentiated, $\hat{\delta}_{SE}$ responds less strongly to an increase in network effects.

In a nutshell, an increase in indirect network externalities has two opposing effects on two-sided collusion. First, if one side values members on the other market side more highly, Nash prices on the latter side fall because competition for this side gets harsher. As a consequence, punishment profits are a falling function of growing indirect network effects. In addition, consumers' utility from platform participation increases if they enjoy a larger benefit from the presence of platform members on the opposing market side. Hence, two-sided platforms can earn larger collusive profits as network externalities grow. Therefore, the collusive gain $(\pi^C - \pi^N)$ is a positive function of indirect network effects. Second and countervailing, however, two-sided platforms also earn larger profits from deviation as network effects become stronger - a result which is due to self-enforcing demand reactions in presence of interrelating network effects. Comparing these opposing effects, it is straightforward to see that consumers do not enjoy additional network benefits when platforms move from Nash competition to collusive prices because market shares stay at $1/2$ in both scenarios. Thus, platforms' gains from collusion follow solely from higher monopoly prices. In case of defection, however, market shares of the defector increase above $1/2$ on both sides. Therefore, consumers receive a higher network benefit from joining the defecting platform. The deviator exploits this effect by lowering his prices just enough to induce a very profitable feedback loop. Solving for the critical discount factor that balances the gains from collusion and optimal defection, it is thus intuitive that the latter asymmetric defection scenario always dominates the symmetric cartel. Collusion becomes harder as symmetric network externalities a increase.

2.3.1.2 Asymmetric Externalities

Having in mind the result of the previous subsection, the question is whether collusion will also be harder to sustain as network externalities grow if those externalities are asymmetric. To this end, I assume from now on that $a_1 = a + \Delta$ and $a_2 = a - \Delta$ with $a, \Delta > 0$. In the newspaper example, this would imply that advertisers (on side 1) care more about readers than readers do care about ads. If Δ is large enough, i.e. if $\Delta > a$, it might even be true that readers dislike advertisements.

Given this network effects specification, equations (2.10) and (2.11) yield the optimal defection prices p_1^D and p_2^D as functions of collusive prices p_1^C and p_2^C:[25]

$$p_1^D = \begin{cases} \frac{p_1^C+t}{2} + \frac{(a+\Delta)(p_2^C-p_2^D)}{2t} - \frac{a-\Delta}{t}\left(\frac{a+\Delta+p_2^D}{2}\right) & \text{if } k < \underline{k}_1 \\ p_1^C - t + \frac{(a+\Delta)(a-\Delta)}{t} + \frac{(a+\Delta)}{t}(p_2^C - p_2^D) & \text{if } k \geq \underline{k}_1 \end{cases} \tag{2.15}$$

$$p_2^D = \begin{cases} \frac{p_2^C+t}{2} + \frac{(a-\Delta)(p_1^C-p_1^D)}{2t} - \frac{a+\Delta}{t}\left(\frac{a-\Delta+p_1^D}{2}\right) & \text{if } k < \underline{k}_2 \\ p_2^C - t + \frac{(a+\Delta)(a-\Delta)}{t} + \frac{(a-\Delta)}{t}(p_1^C - p_1^D) & \text{if } k \geq \underline{k}_2 \end{cases} \tag{2.16}$$

$$\text{with } \underline{k}_1 = \frac{7}{2}(t-a) + \frac{\Delta(t-a)}{2(t+a)} \ ; \ \underline{k}_2 = \frac{7}{2}(t-a) - \frac{\Delta(t-a)}{2(t+a)} \ ; \ \underline{k}_2 < \underline{k}_1$$

From equations (2.15) and (2.16), one can easily see that it might be the case that $\underline{k}_2 < k < \underline{k}_1$. Then, the defecting platform finds it optimal to conquer all of market side 2 while still sharing market side 1 with its rival. The opposing case, however, is never optimal. Hence, three different defection scenarios might arise. First, if $k < \underline{k}_2$, the defector will share total demand on both market sides with its competitor. Second, if $k \geq \underline{k}_2$, the deviating platform will set p_2^D to the maximum possible price that still guarantees a deviation demand of one on this market side, i.e. the price which makes the group-2 agent located at the location of its rival indifferent.[26] Given this price p_2^D, the deviation price on side 1 then follows from maximizing the deviation profit $\pi_{ASE}^D = n_1^D p_1^D + p_2^D$ with respect to p_1^D. Thus, $p_1^D = \frac{1}{2}(t - a + 3\Delta + p_1^C)$ and $n_1^D = \frac{1}{4t}(3a - \Delta + p_1^C + t)$ in this second deviation scenario. Finally, for $k \geq \overline{k}_1$,[27] deviation demand on side 1 will also be equal to one. In this third possible scenario, optimal defection leads to full conquest of both market sides by the defector. Summing up, optimal deviation prices and demands are given as follows:

[25]Given $a_1 = a + \Delta$ and $a_2 = a - \Delta$, collusive prices are equal to $p_1^C = k - \frac{t}{2} + \frac{a+\Delta}{2}$ and $p_2^C = k - \frac{t}{2} + \frac{a-\Delta}{2}$.

[26]Note that demand in case of optimal defection results from plugging p_1^D and p_2^D as well as p_1^C and p_2^C into equation (2.3).

[27]The threshold $\overline{k}_1$ follows from setting $n_1^D = \frac{1}{4t}(3a - \Delta + p_1^C + t) = 1$ and solving for k. Hence, $\overline{k}_1 = \frac{7}{2}(t-a) + \frac{\Delta}{2}$.

$$p_1^D = \begin{cases} \frac{1}{4}\left(2k + t - a + \frac{\Delta^2}{t+a} + \frac{2k\Delta}{t-a}\right) & \text{if } k < \underline{k}_2 \\ \frac{2k+t-a+7\Delta}{4} & \text{if } \underline{k}_2 \leq k < \overline{k}_1 \\ \frac{2k-3(t-a-\Delta)}{2} & \text{if } \overline{k}_1 \leq k \end{cases}$$

$$p_2^D = \begin{cases} \frac{1}{4}\left(2k + t - a + \frac{\Delta^2}{t+a} - \frac{2k\Delta}{t-a}\right) & \text{if } k < \underline{k}_2 \\ \frac{4kt-6(t^2-a(a-\Delta))+(a-\Delta)(2k-t+a-\Delta)}{4t} & \text{if } \underline{k}_2 \leq k < \overline{k}_1 \\ \frac{2k-3(t-a+\Delta)}{2} & \text{if } \overline{k}_1 \leq k \end{cases}$$

$$n_1^D = \begin{cases} \frac{1}{8} + \frac{k}{4(t-a)} - \frac{\Delta}{8(t+a)} & \text{if } k < \underline{k}_2 \\ \frac{2k+t+7a-\Delta}{8t} & \text{if } \underline{k}_2 \leq k < \overline{k}_1 \\ 1 & \text{if } \overline{k}_1 \leq k \end{cases}$$

$$n_2^D = \begin{cases} \frac{1}{8} + \frac{k}{4(t-a)} + \frac{\Delta}{8(t+a)} & \text{if } k < \underline{k}_2 \\ 1 & \text{if } \underline{k}_2 \leq k \end{cases}$$

Given these findings, the deviation profit can be easily calculated as a piecewise function depending on the model parameters k, t, a and Δ:

$$\pi_{ASE}^D = \begin{cases} \frac{(2k+t-a)^2}{16(t-a)} + \frac{\Delta^2}{16(t+a)} & \text{if } k < \underline{k}_2 \\ \frac{(2k+7a-\Delta)^2+2t(18k-a+7\Delta)-47t^2}{32t} & \text{if } \underline{k}_2 \leq k < \overline{k}_1 \\ 2k - 3t + 3a & \text{if } \overline{k}_1 \leq k \end{cases}$$

It is easy to show that deviation profits are non-decreasing in Δ and a over the whole range of k. More specifically, $\frac{\partial \pi_{ASE}^D}{\partial \Delta} > 0$ if $k < \overline{k}_1$, while deviation profits do not depend on the asymmetry of network externalities at all if deviation demand is equal to one on both market sides. Furthermore, deviation profits are falling in the differentiation parameter t, whereas their derivative with respect to k is non-negative.[28]

In order to draw a conclusion concerning the total effect on collusive sustainability,

[28] The exact partial derivatives of the deviation profit are relegated to appendix A.1.

it is important to note that collusive and Nash profits did not change compared to the symmetric case. This is due to the fact the the overall externality or total network benefit $(a_1 + a_2)$ stays constant for all Δ. Nash prices, however, are now given by $p_1^N = t - (a - \Delta)$ and $p_2^N = t - (a + \Delta)$, while maximum collusive prices have changed to be equal to $p_1^C = k + \frac{a+\Delta}{2} - \frac{t}{2}$ and $p_2^C = k + \frac{a-\Delta}{2} - \frac{t}{2}$. As a consequence of these constant Nash and collusive profits, a platform's gain from collusion $(\pi^C - \pi^N)$ is not different from the one under symmetric externalities. The possible gain from defection $(\pi_{ASE}^D - \pi^C)$, however, changes if $\Delta > 0$. In order to investigate if asymmetric network externalities make defection more desirable for a given total network benefit $(a_1 + a_2)$, it suffices to analyze the difference in deviation profits $(\pi_{ASE}^D - \pi_{SE}^D)$:

$$\pi_{ASE}^D - \pi_{SE}^D = \begin{cases} \frac{\Delta^2}{16(t+a)} & \text{if } k < \underline{k}_2 \\ \frac{(2k+7a-\Delta)^2-49t^2}{32t} + \frac{7(2k+\Delta)}{16} - \frac{k^2}{4(t-a)} & \text{if } \underline{k}_2 \leq k < \overline{k} \\ \frac{(2k-7(t-a)-\Delta)^2}{32t} & \text{if } \overline{k} \leq k < \overline{k}_1 \\ 0 & \text{if } \overline{k}_1 \leq k \end{cases}$$

Using assumptions 2.1 and 2.2 and taking into account the thresholds implied by the piecewise functions π_{ASE}^D and π_{SE}^D, it is easy to show that the difference in profits is bigger than zero unless both deviation demands equal one. As can be seen in figure 2.2 for given values of t, a and Δ, the difference in profits falls as the intrinsic utility k increases and becomes zero if $k \geq \overline{k}_1$. In addition, $(\pi_{ASE}^D - \pi_{SE}^D)$ is a non-negative function of the asymmetry Δ and non-increasing in the total network benefit a, but first decreases and then increases in t.[29]

Summing up the above analysis, I conclude that the gain from deviation increases as the network externalities between the two market sides become more asymmetric if k is not too large. Hence, for the incentive constraint as stated in (2.6) to be valid, the critical discount factor $\hat{\delta}_{ASE}$ has to be larger than $\hat{\delta}_{SE}$ as long as $k < \overline{k}_1$. For $k \geq \overline{k}_1$, deviation demand on both sides amounts to one and does no longer

[29] The partial derivatives of $(\pi_{ASE}^D - \pi_{SE}^D)$ are given in appendix A.2.

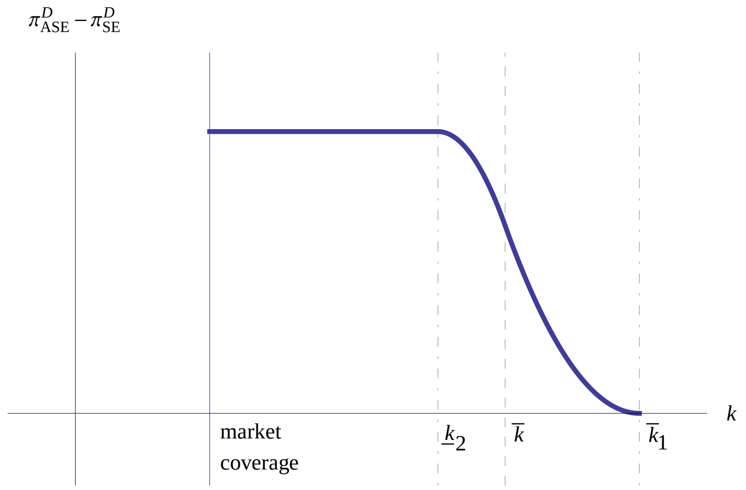


Figure 2.2: difference between deviation profits in case of asymmetric vs. symmetric externalities (for $t = 12$, $a = 4$ and $\Delta = 10$)

depend on Δ. As a consequence, $\hat{\delta}_{ASE}$ is identical to $\hat{\delta}_{SE}$ in that case.

$$\hat{\delta}_{ASE} = \begin{cases} \frac{-9a^3-(\Delta^2+(2k-3t)^2)t+a^2(-12k+9t)+a(\Delta^2-4k^2+9t^2)}{15a^3+a^2(4k-15t)+a(\Delta^2-4k^2-15t^2)+t(-\Delta^2-4k^2-4kt+15t^2)} & \text{if } k < \underline{k}_2 \\ \frac{49a^2+\Delta^2-4k\Delta+4k^2+14\Delta t+4kt-31t^2-2a(7\Delta-14k+9t)}{49a^2+\Delta^2-4k\Delta+4k^2+14\Delta t+36kt-79t^2+a(-14\Delta+28k+30t)} & \text{if } \underline{k}_2 \leq k < \overline{k}_1 \\ \frac{2k-5t+5a}{4k-8t+8a} & \text{if } \overline{k}_1 \leq k \end{cases}$$

As can be seen from above, $\hat{\delta}_{ASE}$ is an intricate parametric function.[30] Therefore, the following figures are provided to shed some light on its behavior with respect to the size and asymmetry of indirect network externalities. For given values of t and k, namely $t = 36$ and $k = 50$, both critical discount factors $\hat{\delta}_{ASE}$ and $\hat{\delta}_{SE}$ are drawn as functions of the total network benefit a. In all three graphs, the leftmost dashed vertical line corresponds to $k = \underline{k}_2$, the one in the middle corresponds to $k = \overline{k}$, and the rightmost indicates where $k = \overline{k}_1$. Going from figure 2.3 to 2.5, the asymmetry between network externalities a_1 and a_2 is increased.

The first thing to note is that both discount factors increase in a over their respective domain ranging from the minimum total benefit needed to fulfill assumption 2.2 up to the maximum threshold for a implied by assumption 2.1, i.e. $a < t$.[31]

[30]The derivatives of $\hat{\delta}_{ASE}$ with respect to the model parameters are omitted for the sake of shortness.

[31]The red vertical line indicates the minimum level of a needed for assumption 2.2 to be fulfilled

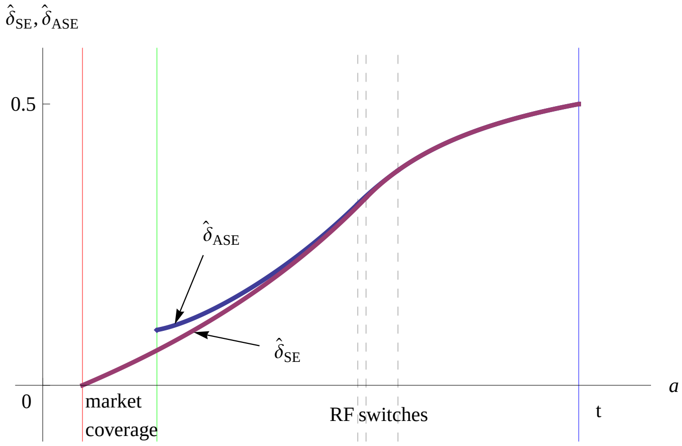


Figure 2.3: critical discount factor for asymmetric and symmetric externalities ($\Delta = 15$)

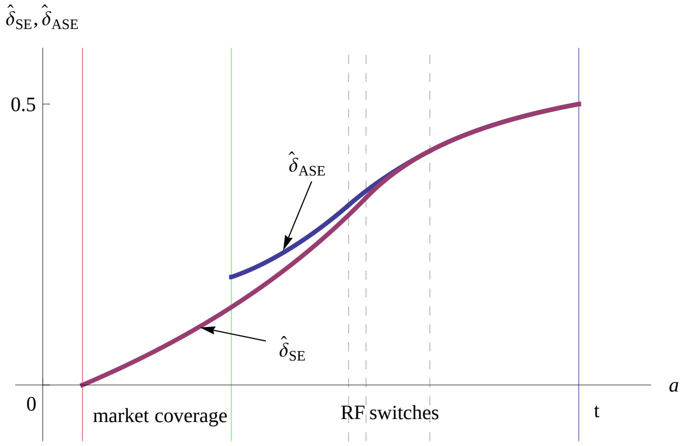


Figure 2.4: critical discount factor for asymmetric and symmetric externalities ($\Delta = 30$)

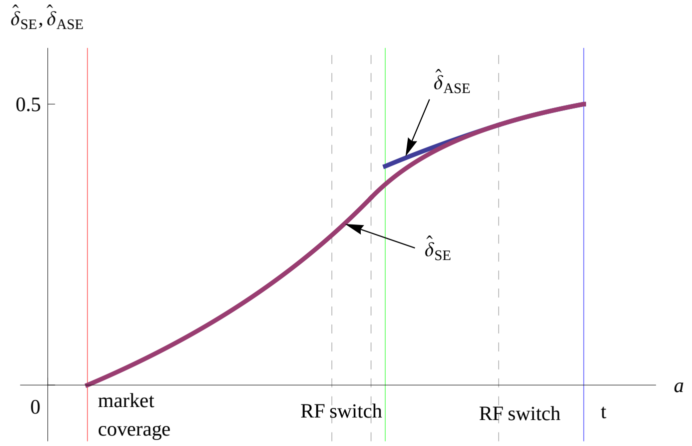


Figure 2.5: critical discount factor for asymmetric and symmetric externalities ($\Delta = 60$)

Furthermore, both functions completely overlap to the right of the last dashed line (i.e. $k = \overline{k}_1$) and converge to $1/2$ for $a \to t$. In this segment of $\hat{\delta}_{ASE}$, optimal defection leads to full market conquest on both sides and the difference in deviation prices due to Δ exactly cancels out. Hence, the asymmetry no longer creates a higher incentive to deviate. To the left of the above mentioned dashed line, $\hat{\delta}_{SE}$ is always smaller than $\hat{\delta}_{ASE}$, and the size of the difference for a given total benefit a increases in Δ as one moves from figure 2.3 to 2.5. A last detail worth mentioning is that the domain of $\hat{\delta}_{ASE}$ narrows down as Δ grows. In consequence, it might be the case that optimal defection implies market conquest of one or even both sides over the whole range of a that fulfill assumptions 2.1 and 2.2 for given values of k and t.

To conclude this section, the following proposition summarizes the effect of asymmetric externalities on the sustainability of a two-sided cartel.

Proposition 2.2 *Suppose that assumptions 2.1 and 2.2 are fulfilled and externality parameters can be described as $a_1 = a + \Delta$ and $a_2 = a - \Delta$. Then, the following holds true:*

- $\frac{\partial(\pi^D_{ASE} - \pi^N)}{\partial \Delta} \geq 0$ *and* $\frac{\partial(\pi^C - \pi^N)}{\partial \Delta} = 0$, *i.e. the incentive to deviate from the collusive agreement increases while the gains from collusion stay constant as*

in case of $\Delta = 0$ while the green vertical line marks the minimum level of a for assumption 2.2 to hold true in case of $\Delta > 0$ as given in the subtitle of the respective figure.

Δ gets larger. Thus, $\frac{\partial \hat{\delta}_{ASE}}{\partial \Delta} \geq 0$.

- *The size of this asymmetry effect is decreasing in the total network benefit and the intrinsic utility parameter: $\frac{\partial^2 \hat{\delta}_{ASE}}{\partial \Delta \partial a} \leq 0$, $\frac{\partial^2 \hat{\delta}_{ASE}}{\partial \Delta \partial k} \leq 0$.*

- *For $k \geq \overline{k}_1$, $\frac{\partial \hat{\delta}_{ASE}}{\partial \Delta} = 0$, because $n_1^D = n_2^D = 1$ and the asymmetries between the two sides exactly balance each other out.*

- *The critical discount factor increases in the total network benefit $(a_1 + a_2) = a$: $\frac{\partial \hat{\delta}_{ASE}}{\partial a} > 0$.*

Starting out with symmetric network externalities in 2.3.1.1, I have shown that collusion becomes harder to sustain as these network effects increase. If one now allows for an increasing asymmetry between both sides for a given total network benefit a, this will have a negative impact on collusive sustainability if the intrinsic utility from joining a platform is not too high. The driving force for this result is that the external benefits of a price decrease are smaller when network externalities are asymmetric rather than symmetric. In detail, the deviation price on side 1 increases by a larger amount than the deviation price on side 2 falls after an increase of the network benefit parameter a_1 by Δ and an equally sized reduction in a_2. In contrast, demand reactions to Δ are perfectly symmetric if $k < \underline{k}_2$. Deviation demand on side 1 falls while side-2-demand increases by the same amount. For $\underline{k}_2 \leq k$, deviation demand on side 2 is equal to one as all consumers on this side are sufficiently subsidized to overcome their transportation costs. Deviation demand on side 1, however, is still smaller than one as long as $k < \overline{k}_1$. In consequence, the external benefit of a price decrease becomes slightly smaller up until $k = \overline{k}_1$, the point where the asymmetry no longer biases the incentive to deviate because both market sides are completely overtaken by the defector. In sum, deviation profits in case of asymmetric externalities are larger than or equal to deviation profits in presence of symmetric network effects while collusive and Nash profits do not change in Δ.

The results presented in 2.3.1.1 and 2.3.1.2 are in line with Evans & Schmalensee's hypothesis that collusion is harder to sustain in two-sided markets. In contrast to their argumentation, however, my findings do not hinge on increased monitoring or coordination costs. The bottom line that collusion becomes harder as indirect network effects grow rather follows from balancing two opposing effects. While it might be due to the chosen linear Hotelling framework that the deviation effect always dominates, I would like to emphasize that the existence of these two countervailing effects caused by network externalities in a two-sided market, namely higher gains from colluding as well as larger benefits from defection, is of a more general nature.

Another way to interpret my results would be to compare the impact of indirect network effects in my framework to a situation where platforms are perfectly compatible.[32] In the latter case, users on one market side enjoy a network benefit from all consumers on the other market side independent of their platform choice. In consequence, a platform's market share on one side of the market does not influence consumers' decision making on the other - the only two things that matter are a consumer's location on the respective Hotelling line and the price he faces on his market side. A platform's incentive to deviate from the collusive agreement is much smaller in case of such a compatibility because of the lacking feedback effects. Thus, collusion will always be harder to sustain in case of no compatibility whereas the scope for collusion is larger when platforms do not have to compete for indirect network externalities.

2.3.1.3 Extension: One-Sided Pricing in a Two-Sided Market

To complete the section on two-sided collusion, let me check if my results carry over if platforms can only charge positive prices to one group of customers, as it is the case for the free-to-air TV channels ProSiebenSAT1 and RTLGroup mentioned in the introduction. More generally speaking, I investigate whether collusion is harder to sustain as network effects increase in those two-sided markets where platforms have to earn all their money from advertisers, as it is also true for free newspapers or yellow pages. In these industries, advertisers gain utility from a large viewer- or readership. Thus, their network parameter is positive and they are willing to pay a positive price for accessing readers. Readers or viewers, on the other hand,

[32]See Katz and Shapiro (1985) for a seminal contribution on network externalities and compatibility.

very often dislike ads in these markets.[33] As a consequence, newspapers or TV channels would sometimes like to subsidize them in order to attract enough readers or viewers to maximize profits gained from advertisers. Negative prices, however, are seldom feasible. Thus, if prices on the viewers' market side are fixed to be zero, this might add an additional twist to the effect of indirect network externalities on collusive sustainability.

I will shortly analyze the collusive game described in section 2.2 under the additional assumption that w.l.o.g. $a_1 > 0$, $a_2 < 0$ and $p_2^A = p_2^B = 0$.[34]

Given these assumptions, demand simplifies to:

$$n_1^i = \frac{1}{2} + \frac{t(p_1^j - p_1^i)}{2(t^2 - a_1 a_2)} \ ; \ n_2^i = \frac{1}{2} + \frac{a_2(p_1^j - p_1^i)}{2(t^2 - a_1 a_2)} \qquad i, j \in \{A, B\} \ ; \ i \neq j \tag{2.17}$$

Note that demand on both sides is influenced by the price difference on side 1. While platforms can earn profits only on the advertiser side 1, they have to take into account the positive external benefit an additional consumer on side 2 imposes on side-1 consumers when making their pricing decision. Nash prices in this special case of the framework are equal to $p_1^N = t - \frac{a_1 a_2}{t}$ and profits amount to $\pi^N = \frac{t}{2} - \frac{a_1 a_2}{2t}$. Hence, in case of zero prices on side 2, the external benefit simply amounts to the additional benefit a_1 that each side-1 agent enjoys from one more agent on the other side, i.e. the extra revenue a TV channel can extract from its advertisers when attracting one more viewer. This amount is multiplied by the fraction a_2/t of viewers who are lost by one more ad displayed on TV. Thus, the price that advertisers have to pay increases the more viewers dislike seeing advertisements.

Collusive profits follow from symmetric market sharing and setting p_1^C equal to $k - \frac{t}{2} + \frac{a_1}{2}$ and thereby extracting all utility from the indifferent advertiser on side 1 and are given by $\pi^C = \frac{k}{2} - \frac{t}{4} + \frac{a_1}{4}$. Thus, collusive profits increase if advertisers value readers more highly, as it was the case in the previous subsection. The effect of a

[33]See, for example, Wilbur (2008) who finds that consumers are strongly advertising averse in the television industry.

[34]Note that Armstrong and Wright (2007) also analyze the case when non-negative prices are not allowed and find Nash prices similar to the ones presented below. They assume, however, that network externalities are both positive.

higher viewership valuation on Nash profits, however, differs from the general case. For a given negative externality that ads impose on viewers, the Nash price that a newspaper or TV channel demands from advertisers rises in a_1. As a consequence, the gain from colluding only increases in advertisers' benefits from viewers a_1 if viewers do not dislike ads too much. On the other hand, the gain from colluding always decreases if viewers object to ads very strongly. In this case, TV channels can demand a higher competitive ad price while collusive prices stay constant.

Turning to the incentive to deviate from the collusive agreement, deviation prices can be inferred from a platform's optimal reaction function which is given as follows if $a_2 > -t$:[35]

$$p_1^D = \begin{cases} \frac{p_1^C+t}{2} - \frac{a_1a_2}{2t} & \text{if } p_1^C < \frac{3(t^2-a_1a_2)}{t} \\ p_1^C - t + \frac{a_1a_2}{t} & \text{if } p_1^C \geq \frac{3(t^2-a_1a_2)}{t} \end{cases} \tag{2.18}$$

Plugging both collusive and defection prices into equation (2.18) yields deviation profits:

$$\pi^D = \begin{cases} \frac{(a_1(t-2a_2)+t(2k+t))^2}{32t(t^2-a_1a_2)} & \text{if } k < \overline{\overline{k}} \\ k + \frac{a_1}{2} - \frac{3}{2}t + \frac{a_1a_2}{t} & \text{if } k \geq \overline{\overline{k}} \end{cases} \tag{2.19}$$

$$\text{with } \overline{\overline{k}} = \frac{7}{2}t - \frac{a_1(t+6a_2)}{2t}$$

It is easy to show that the deviation price is non-monotonic in both network effects. Moreover, deviation demand on side 1 increases in the difference between collusive and defection prices. Due to the negative network benefits of viewers, however, the more advertisers are attracted by the defector the less viewers are willing to stay with her. As a consequence, the gains from defection $(\pi^D - \pi^C)$ fall in the absolute

[35]Note that if the negative externality of readers is too large, i.e. if $a_2 \leq -t$, then deviation demand on side 2 will be equal to zero at some point. If this is the case, platforms' optimal defection would no longer depend on a_2. Thus, I focus on the situation where readers do not dislike ads too much. If $a_2 > -t$, deviation demand on side 1 increases faster than deviation demand on side 2 falls as the defector lowers its price and attracts more advertisers. Thus, it will be the case that $n_1^D = 1$ if a_1 is large enough. At that point, the optimal defection price is chosen such that the defector just covers the market on side 1. Given $n_1^D = 1$, deviation demand on side 2 does no longer change because viewers already suffer from the maximum number of ads. In fact, side-2 deviation demand will then only converge to zero if $a_2 \to -t$.

value of a_2. The impact of advertisers' benefits from viewers a_1 on gains from collusion, on the other hand, is only positive if the negative externalities a_2 that viewers have to incur are not too large in absolute value. Summing up the previous findings on the incentives to collude or to deviate, I expect the critical discount factor $\hat{\delta}_{PF}$ to be non-monotonic in the network externality a_1, too. In other words, if one price is fixed to be equal to zero, it might be the case that collusion becomes easier to sustain as network externalities become stronger.

Noticing that $\hat{\delta}_{PF}$ is given by the following equation,

$$\hat{\delta}_{PF} = \begin{cases} \frac{2k - 3t + a_1 + 2\frac{a_1 a_2}{t}}{2k + 5t + a_1 - 6\frac{a_1 a_2}{t}} & \text{if } k < \overline{\overline{k}} \\ \frac{2k - 5t + a_1 + 4\frac{a_1 a_2}{t}}{4k - 8t + 2a_1 + 6\frac{a_1 a_2}{t}} & \text{if } k \geq \overline{\overline{k}} \end{cases} \tag{2.20}$$

it is easy to see that $\hat{\delta}_{PF}$ increases monotonically in a_2, while being non-monotonic in a_1. If $a_2 < t^2/(t-2k)$, then $\hat{\delta}_{PF}$ falls in a_1, whereas it increases in a_1 for $a_2 > t^2/(t-2k)$. Put differently, the less viewers dislike ads, the harder it is for TV channels to collude. Further, if viewers' distaste of ads is not too high, then collusion is less likely when advertisers gain a larger utility from their readership. If, however, viewers hate to watch ads very much, collusion actually becomes easier when advertisers' valuation of viewers increases, because the negative impact of readers' distaste for ads on gains from deviation is larger than its negative influence on gains from collusion. The following remark summarizes these findings:

Lemma 2.3 *If $a_1 > 0$, $-t < a_2 < 0$ and $p_2^A = p_2^B = 0$, collusion might become easier to sustain as a_1 grows as long as the negative network externality a_2 is large enough in absolute value, i.e. $\frac{\partial \hat{\delta}_{PF}}{\partial a_1} > 0$ if $a_2 > t^2/(t-2k)$ and $\frac{\partial \hat{\delta}_{PF}}{\partial a_1} < 0$ if $a_2 < t^2/(t-2k)$.*

The intuition for this result, which differs from the one presented in propositions 1 and 2, hinges on the fact that viewers cannot be priced. Thus, newspapers or free-to-air TV channels have to earn all their money from advertisers. When viewers dislike ads, they will prefer to watch a channel that shows fewer ads. Hence, as a_1 increases, TV channels might prefer to collude and share the number of advertisers, rather than to deviate and loose viewers. The result, however, is driven by the fact that full market coverage is assumed for both market sides, i.e. viewers always watch one of the two channels no matter how much they dislike the advertisements they will have to face. This could be rationalized by the fact that viewers strongly like both TV channels for their program, i.e. for k, such that they still prefer watching TV over not watching.

2.3.2 One-sided Collusion

Some of the cartel cases described in the introduction imply that platforms could possibly find it more profitable to collude only on one market side. Consider, for example, the case of German nationwide newspapers that planned to form a common agency for selling job ad space. They decided to collude on prices on the advertiser side. The newspaper cartels detected in Switzerland and Australia, on the other hand, were fixing cover prices. One should therefore ask if collusion only on one side of the market might actually be easier to implement or yield higher profits than full collusion on both sides under certain conditions. Furthermore, analyzing one-sided collusion allows for an evaluation of Evans and Schmalensee (2008)'s claim that all supra-competitive profits earned on the cartelized market side will be competed away on the other and therefore one-sided collusion will not be profitable at all.

2.3.2.1 Symmetric Externalities

Let me start by analyzing the symmetric case of $a_1 = a_2 = a > 0$ and assuming w.l.o.g. that platforms collude on side 1. If firms collude on the highest possible price given that markets are equally split, they will set $p_1^{OC} = k + a/2 - t/2$. On side 2, they will now maximize profits separately taking into account p_1^{OC}, which yields prices equal to $p_2^{OC} = t - \frac{a}{t}(a + p_1^{OC})$. Recalling assumption 2.2, it is easy to show that p_2^{OC} is always smaller than $p_2^N = t - a$. Thus, platforms do compete away some of the supra-competitive profits earned on side 1 as Evans and Schmalensee (2008) expected it to be the case.[36] Actually, the collusive profit amounts to $\pi^{OC} = \frac{(t-a)}{2t}(t + a + p_1^{OC}) = \frac{(t-a)}{4t}(2k + 3a + t)$, which is always smaller than the collusive profit when platforms collude on both sides. Nevertheless, it is still beneficial to collude for all feasible parameter values. In contrast to section 2.3.1.1, however, the gain from colluding is no longer monotonically increasing in a, but instead falls for $k > 3(t - a)$.

Lemma 2.4 *If $a_1 = a_2 = a > 0$ and platforms collude only on side 1, some of the supra-competitive profits earned on the colluding market side will be competed away by setting prices below p_2^N on the opposite side. Yet, one-sided collusion is always profitable, i.e. $(\pi^{OC} - \pi^N) \geq 0$, but $\frac{\partial(\pi^{OC} - \pi^N)}{\partial a} \geq 0$ only if $k \leq 3(t - a)$.*

Turning to a platform's incentive to deviate, it is important to note that it is never an optimal reaction for the defecting platform to fully conquer market side 2, where firms are already competing. Instead, optimal defection prices on side 2 are now actually higher than p_2^{OC}. While it is optimal to decrease prices on market side 1, which in return raises market shares on this side, the deviation price on side 2 is increased by such an amount that the corresponding market share of the defecting platform stays constant at 1/2. Hence, reaction functions and deviation profits are given as follows:

$$R_1^D(p_1^{OC}) = \begin{cases} \frac{p_1^{OC}+t}{2} - \frac{a}{2} & \text{if } k < \tilde{k} \\ \frac{(t^2-a^2)}{t^2}(p_1^{OC} - t + a) & \text{if } k \geq \tilde{k} \end{cases} \quad \text{with } \tilde{k} = \frac{7}{2}t - \frac{3}{2}a$$

$$R_2^D(p_2^{OC}) = \frac{p_2^{OC} + t}{2} - \frac{a}{2} \qquad \forall k$$

$$\pi_{OC}^D = \begin{cases} \dfrac{15(t^2 - a^2) + 2t(t - a) + 4k(k + t - a)}{32t} & \text{if } k < \tilde{k} \\ \dfrac{4kt - 4t^2 - 2ka + 5at - 3a^2}{4t} & \text{if } k \geq \tilde{k} \end{cases} \tag{2.21}$$

[36]This finding is also in line with Armstrong and Wright (2007). They shortly discuss the implications of inflated prices and conclude that platforms gain relatively little from price fixing on a single side (see p. 360).

One can easily show that the gain from deviation ($\pi^D_{OC} - \pi^{OC}$) is an increasing function of a. Hence, the larger the network externalities a become, the more profitable it is for a platform to deviate from the one-sided collusive agreement. Moreover, comparing optimal defection strategies with those under two-sided collusion (see equation (2.12)), it turns out that price reactions are identical if $k < \tilde{k}$. Yet, collusive prices differ under the two different collusive practices. In addition, the defection price on side 1 will be a less strong, but still positive function of the collusive price on side 1 when $k \geq \tilde{k}$, i.e. when the defector attracts all consumers on side 1.

Balancing the gain from deviation and the incentive to collude gives me the critical discount factor $\hat{\delta}_{OC}$ above which monopoly prices on market side 1 can be sustained in presence of symmetric externalities:

$$\hat{\delta}_{OC} = \begin{cases} \frac{2k+3a-3t}{2k-5a+5t} & \text{if } k < \tilde{k} \\ \frac{t(2k+3a-5t)}{t(2k+3a-5t)-a(2k+3a-t)} & \text{if } k \geq \tilde{k} \end{cases} \tag{2.22}$$

Comparing this discount factor to the one for two-sided collusion as given in equation (2.14), one sees that they are identical if $k < \overline{k}$. In other words, when network externalities are small compared to transport costs and intrinsic utility levels, it is not harder to sustain collusion on both sides rather than only on side 1. If network externalities become large enough such that $k \geq \overline{k}$, however, it is even harder to sustain collusion only on one side. The intuition for this result follows directly from platforms' incentives to collude. While a firm always benefits more from colluding as a increases in case of price fixing on both market sides, this is not true for one-sided collusion. Instead, a platform earns less and less from colluding on side 1 if consumers enjoy higher network benefits because this enforces price competition on the remaining side.[37] In consequence, more and more of the collusive profits earned on side 1 will be lost on side 2. Gains from deviation, on the other hand, always increase as network benefits become stronger both under one-sided and two-sided collusion.

[37]This is exactly the case when $k > 3(t - a)$.

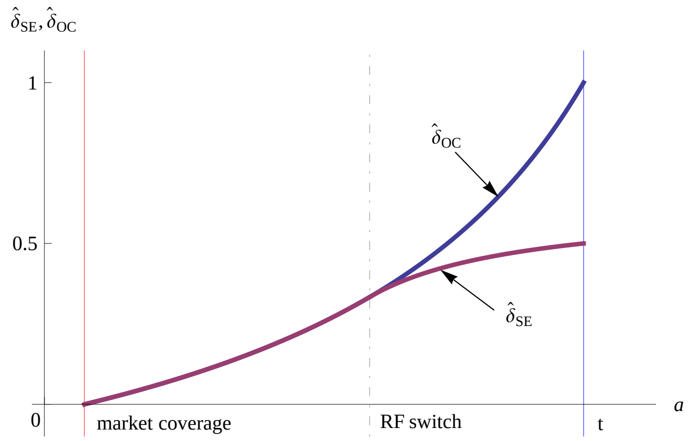


Figure 2.6: critical discount factor for one- and two-sided collusion ($\Delta = 0$)

For $k = 50$ and $t = 36$, critical discount factors $\hat{\delta}_{OC}$ and $\hat{\delta}_{SE}$ are plotted as a function of a in figure 2.6.[38]

The following proposition sums up the above findings for symmetric externalities:

Proposition 2.3 *Suppose that $a_1 = a_2 = a > 0$ and platforms collude only on side 1. Then, one-sided collusion is harder to sustain than collusion on both sides if network effects are large, i.e. $\hat{\delta}_{OC} > \hat{\delta}_{SE}$ if $k > \overline{k}$. If $k < \overline{k}$, $\hat{\delta}_{OC} = \hat{\delta}_{SE}$.*

2.3.2.2 Asymmetric Externalities

In case of asymmetric externalities $a_1 = a + \Delta$ and $a_2 = a - \Delta$, platforms' gains from cartelization depend on which side they decide to collude on. Given that they are free to choose a market side, collusion will take place *on side 2* because the gain from collusion is always greater or equal to zero if platforms cartelize on the side that imposes higher indirect externalities. For media markets, it seems very reasonable to assume that advertisers benefit more from readers than vice versa. Thus, newspapers can earn the higher profits when colluding on cover prices instead of ad prices. It is therefore unsurprising that the newspaper cartels in Australia and Switzerland did so.

Coming back to my model, I find that platforms will set $p_2^{OC2} = k - \frac{t}{2} + \frac{a-\Delta}{2}$ on the colluding market side 2 and $p_1^{OC2} = t - \frac{(a-\Delta)}{t}(a + \Delta + p_2^{OC2})$ on the competitive

[38]The label "RF switch" indicates the value of a at which $k = \overline{k}$.

side 1, which results in the following collusive profit:

$$\pi^{OC2} = \frac{1}{2}p_1^{OC2} + \frac{1}{2}p_2^{OC2} = \frac{(t-a+\Delta)(2k-t+a-\Delta)}{4t} + \frac{t^2-a^2+\Delta^2}{2t}$$

Note that the optimal reaction on the still competitive side 1 will lead to a price p_1^{OC2} below Nash only if the asymmetry is relatively small compared to the overall network benefit, namely if $\Delta < a$. This means that both network effects a_1 and a_2 need to be positive in order to have fiercer price competition on the non-colluding side 1. In other words, platforms will only compete away parts of the supra-competitive profits to be earned on side 2 if $\Delta < a$. If $\Delta \geq a$, however, side-2 agents actually dislike side-1 agents. Yet, platforms still want them to join because they know that these side-2 agents impose large positive network benefits on the other group. Hence, in case of Nash competition, consumers on side 2 will be subsidized by paying price $p_2^N < t$. Side-1 agents, on the other hand, are targeted much less aggressively and pay a price $p_1^N > t$ for having access to side 2. In case of collusion on side 2, platforms can extract all intrinsic utility from side-2 agents. They have to diminish the collusive price, though, by transport costs and the negative externality a_2 that the indifferent consumer at $1/2$ receives from type-1 agents joining the platform. By assumption 2.2, this collusive price p_2^{OC2} is always larger than the Nash price. Hence, platforms make supra-competitive profits on side 2. On side 1, the optimal price reaction is a positive function of collusive prices p_2^{OC2}.[39] Thus, competition for side-1 consumers is actually attenuated. In sum, collusion only on side 2 might therefore lead to higher collusive profits than two-sided collusion. In fact, for $\Delta \geq t+a$, the collusive profit π^{OC2} will be bigger than π^C (see equation (2.9)) for all valid ranges of k.

Turning a platforms gain from colluding on side 2 only,

$$\pi^{OC2} - \pi^N = \frac{(t-a+\Delta)(2k-3(t-a)+\Delta)}{4t}$$

I find that it is always profitable to collude just on market side 2. Therefore, the following lemma can be stated.

[39]It is easy to see from equation (2.10) that $\frac{\partial R_1^i}{\partial p_2^i} > 0$ if $a_2 < 0$.

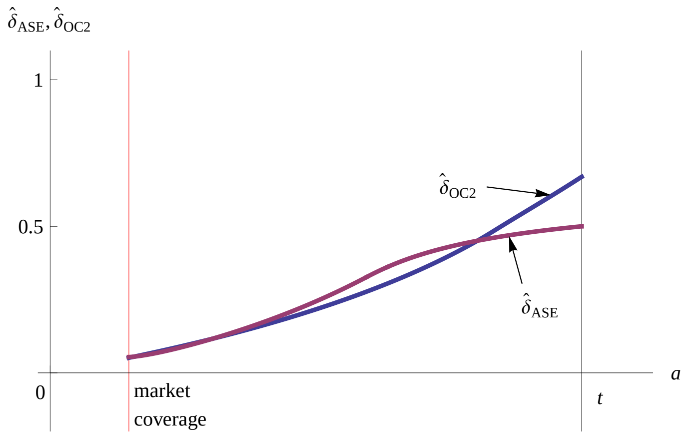


Figure 2.7: One-sided collusion on side 2 ($\Delta = 8$)

Lemma 2.5 *If $a_1 = a+\Delta$ and $a_2 = a-\Delta$, platforms prefer colluding on side 2 over colluding on side 1. Collusion only on side 2 is always profitable, i.e. $(\pi^{OC2} - \pi^N) \geq 0$. Moreover, if a_2 is strongly negative, i.e. if $\Delta \geq t + a$, then $\pi^{OC2} > \pi^C$.*

Turning to a platform's optimal defection strategy, one has to note that deviation demands do not balance each other out as it was the case in section 2.3.1.2. Instead, deviation demand on side 1 increases faster with falling deviation prices if $\Delta \geq t+a$, while deviation demand on side 2 does so if $\Delta < t + a$. In consequence, prices, demand functions and profits on both market sides in case of optimal defection are quite intricate and therefore omitted for the sake of shortness at this point.[40] The critical discount factor $\hat{\delta}_{OC2}$ above which collusion on side 2 is sustainable is given in appendix A.3 and depicted as a function of a for different values of Δ in figures 2.7 and 2.8.[41]

As one can see from these figures, the critical discount factor $\hat{\delta}_{OC2}$ is increasing in the total network benefit a. Its behavior with respect to the asymmetry between both market sides, however, is not as clearcut. For small a, it increases in Δ, whereas it decreases for larger a. Overall, $\hat{\delta}_{OC2}$ might be bigger or smaller than the critical discount factor for two-sided collusion $\hat{\delta}_{ASE}$ depending on the size of the model parameters. Hence, one cannot draw the same conclusion as for symmetric

[40] The Mathematica notebook containing the respective formulas is available from the author upon request.

[41] The chosen numerical values for k and t are 50 and 36, respectively.

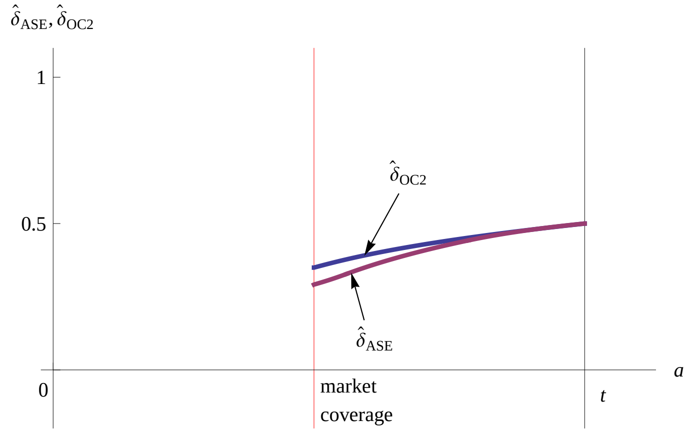


Figure 2.8: One-sided collusion on side 2 ($\Delta = 45$)

externalities, i.e. one-sided collusion is not necessarily harder to sustain than two-sided collusion.

To conclude this subsection, let me consider the case where collusion on side 2 is not possible, e.g. because of prohibitively high monitoring costs or a severe risk of revelation by antitrust authorities. This was exactly the case for the three German newspapers that wanted to collude on job advertisement prices. They thought that such a form of coordinated behavior would be tolerated by the German antitrust authority while collusion on cover prices would never have been accepted. Yet, these newspapers must have believed that such a cooperation on the ad prices would allow them to increase profits.

To shed some light on the incentives to collude on the less attractive market side, I will analyze the situation where platforms choose to form a cartel on side 1 although collusive gains will be smaller. In fact, colluding on side 1 by setting $p_1^{OC1} = k - \frac{t}{2} + \frac{a+\Delta}{2}$ leads to a price of $p_2^{OC1} = t - \frac{(a+\Delta)}{t}(a - \Delta + p_1^{OC1})$ on side 2. Therefore, collusion on side 1 suffers from the fact that supra-competitive profits earned on side 1 are (partially) competed away by setting prices below p_2^N on side 2. In consequence, one-sided collusion in this form is only profitable as long as $\Delta < t - a$, as one can easily see from the following equation:

$$\pi^{OC1} - \pi^N = \frac{(t - a - \Delta)(2k - 3(t - a) - \Delta)}{4t}$$

This means that only if platforms can demand non-negative Nash prices on side 2, i.e. only if the positive network benefits that consumers on side 1 enjoy are not too large, collusion will be profitable. Thus, in case of an asymmetric network effects structure and collusion on side 1, Evans and Schmalensee's (2008) hypothesis that all supra-competitive profits will be competed away on the non-colluding market side is true for $\Delta \geq t - a$. In other words, if the asymmetry is very large and side-1 customers value the opposing market side much more than vice versa, platforms will prefer not to collude at all. Furthermore, the gain from colluding increases in a only if $k \leq \frac{3}{2}(t-a) - \Delta$ and falls in a for $k > \frac{3}{2}(t-a) - \Delta$. Hence, the non-monotonicity property of the collusive gain with respect to a that was found for symmetric network effects in subsection 2.3.2.1 still holds if platforms cartelize on side 1.

Lemma 2.6 *If $a_1 = a + \Delta$ and $a_2 = a - \Delta$ and platforms collude only on side 1, some of the supra-competitive profits will be competed away by setting $p_2^{OC1} < p_2^N$. Thus, it is only profitable to collude on side 1 if $\Delta < t - a$.*

Suppose from now on that $\Delta < t - a$. Turning to a platform's incentive to deviate, it is never optimal to fully conquer the market on the non-colluding side 2. It might, however, be optimal to conquer market side 1 when the collusive price is high enough. If this is the case, the defection price on side 1 will be set such that demand stays equal to 1 and the optimal reaction on side 2 is adapted accordingly. In consequence, it might even be the case that deviation demand on side 2 falls to zero as collusive prices increase and the defecting platform chooses to only serve consumers on side 1. This result, however, is driven by the fact that both sides always have an intrinsic motivation k to join one of the platforms even in the absence of any consumers on the opposite market side. A detailed analysis of deviation strategies and profits is not displayed at this point, but the resulting critical discount factor that allows for one-sided price collusion on side 1 is displayed in appendix A.4. To exemplify its behavior with respect to both the total benefit a and the asymmetry parameter Δ, the critical discount factor $\hat{\delta}_{OC1}$ is depicted in figures 2.9 and 2.10 for $k = 50$ and $t = 36$.[42]

[42]The blue line indicates the maximum value of a for which the discount factor is still well defined given that Δ needs to be smaller than $(t - a)$ for collusion on side 1 to be profitable.

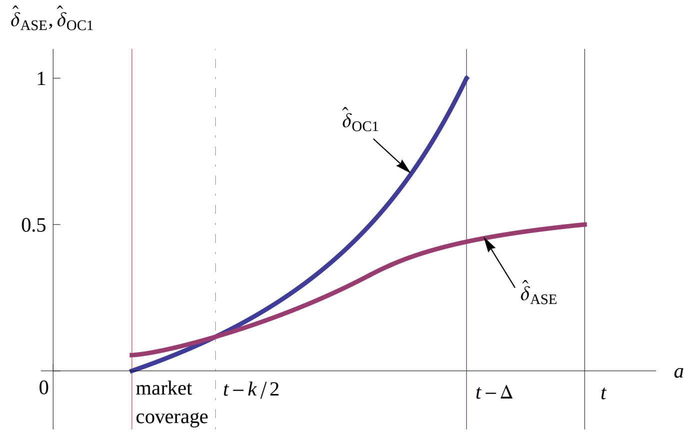


Figure 2.9: One-sided collusion on side 1 ($\Delta = 8$)

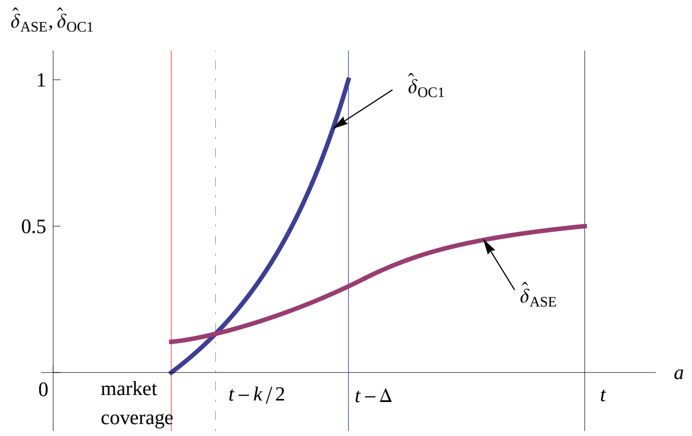


Figure 2.10: One-sided collusion on side 1 ($\Delta = 16$)

Three things are worth noticing in these figures: First, $\hat{\delta}_{OC1}$ increases in a. Second, it is always equal to zero at the lowest possible value of a that still satisfies the market coverage assumption for given values k, t and Δ and rises up to 1 for the maximum level of a under which this one-sided collusion is still profitable. Finally, $\hat{\delta}_{OC1}$ first decreases and then increases in the asymmetry Δ. As a result, collusion only on side 1 is easier to stabilize than two-sided collusion as long as $k < 2(t-a)$ while it is harder to sustain for $k > 2(t-a)$. Hence, colluding only on ad prices is easier for newspapers than colluding on both sides if the intrinsic motivation is small, e.g. if readers' valuation for the content is low.

The following proposition sums up the above sustainability analysis of one-sided price collusion in case of asymmetric externalities:

Proposition 2.4 *Suppose that* $a_1 = a + \Delta$ *and* $a_2 = a - \Delta$ *and platforms can only collude on one of the two sides of the market. Then,* $\frac{\partial \hat{\delta}_{OC1}}{\partial a} \geq 0$ *and* $\frac{\partial \hat{\delta}_{OC2}}{\partial a} \geq 0$. *The comparison of critical discount factors for one-sided and two-sided collusion does not lead to clearcut results:*

- $\hat{\delta}_{OC1} < \hat{\delta}_{ASE}$ *if* $k < 2(t-a)$ *and* $\hat{\delta}_{OC1} \geq \hat{\delta}_{ASE}$ *if* $k \geq 2(t-a)$
- $\hat{\delta}_{OC2} \lesseqgtr \hat{\delta}_{ASE}$ *depending on parameter values of* a, t *and* k

Ultimately, section 2.3.2 refutes Evans & Schmalensee's hypothesis that one-sided collusion is never optimal. Although they have been partially right in claiming that supra-competitive profits earned on the colluding market side will be competed away on the other, this effect is not strong enough to make one-sided cartelization unprofitable in general. Instead, platforms might even earn higher additional profits under one-sided collusion than they do in case of a two-sided cartel if network effects on the colluding side are strongly negative.

Finally, one-sided collusion has different welfare implications than collusion on both sides of the market. First, platforms extract all utility from the indifferent consumer on the colluding market side. Hence, consumer surplus on this side decreases. Market participants on the non-colluding side, however, might benefit from the collusive agreement via a reduction in their membership fees because of the existing externalities. In consequence, the distributional effects of collusion differ: consumers on the colluding side suffer whereas those on the non-colluding side might benefit from cartelization. Overall, total welfare might therefore actually increase, as it has been shown in a different media markets model by Dewenter, Haucap, and Wenzel (2011).

2.4 Multi-Homing

In order to test the robustness of the above results with respect to the form of competition between platforms, I will now allow for consumers to multi-home. Furthermore, the multi-homing setup might be seen as being better suited for some of my motivating examples since there is empirical evidence that advertisers place their ads in several newspapers or TV channels.[43]

In line with the literature, I will assume that those agents that enjoy higher network benefits can choose to multi-home, i.e. given that $a_1 = a + \Delta$ and $a_2 = a - \Delta$, *group-1 consumers* are able to join both platforms.[44] In this case, a platform's Nash profit that will serve as a punishment strategy in case of defection from the collusive agreement results from maximizing stage game profits $\pi_i = p_1^i n_1^i + p_2^i n_2^i$ with respect to prices for each platform i and solving simultaneously. While side-2 demand is still described by equations (2.4) and (2.5), side-1 demand is equal to equation (2.3). Thus, I obtain the following Nash prices:

$$p_1^N = \frac{k}{2} + \frac{a_1 - a_2}{4} \; ; \; p_2^N = t - \frac{a_1}{t}(a_2 + p_1^N) \tag{2.23}$$

Given these symmetric prices, full market coverage on side 2 will be guaranteed if the indifferent consumer located at 1/2 has a non-negative utility from buying. Hence, assumption 2.2 has to be re-adjusted:[45]

[43]Still, one has to take into account that timing plays an important rule when defining whether one side actually multi-homes. In case of two newspapers, would this mean that advertisers need to place their ads in both papers at the same day or would it be enough if they switch between papers from week to week?

[44]Armstrong and Wright (2007) actually derive conditions on the relative sizes of transport costs and network externalities for which one side prefers to multi-home. For details, see p. 356ff.

[45]Note that if $a_1 = a + \Delta$ and $a_2 = a - \Delta$, assumption 2.3 simplifies to $k \geq \frac{3}{2}(t - a) + \frac{\Delta^2}{2(t+a)}$.

Assumption 2.3 $k \geq \frac{6t^2-(a_1^2+4a_1a_2+a_2^2)}{4t+2(a_1+a_2)}$

From equation (2.23), one can see that Nash prices on the single-homing side 2 are identical to the ones shown in equation (2.7). Prices on the multi-homing side, however, neither depend upon the prices set on the opposing market side, nor are they influenced by the strength of differentiation between the two platforms, i.e. the transport cost t. Hence, platforms behave as if they are not competing for type-1 consumers, but instead compete indirectly by attracting side-2 consumers. Thus, I reproduce the qualitative findings of Armstrong and Wright (2007).

Given the above Nash prices, equilibrium demand on side 1 results to be elastic and equal to $n_1^N = (2k+a_1+a_2)/4t$, while $n_2^N = 1/2$. The demand that a platform faces on the multi-homing side in case of competition is actually rather intuitive, i.e. it is increasing in the utility that a group-1 agent enjoys when buying from a certain platform, namely in k and a_1, while it decreases both in the price he has to pay and the transportation costs. Hence, consumers located in the middle between the two platforms will actually multi-home if transport cost are not too high compared to the utility to be gained when joining the second platform. More precisely, the number of multi-homing side-2 agents will be positive if $t < k + (a_1 + a_2)/2$. Nash profits are given by:

$$\begin{aligned}\pi_{MH}^N &= \frac{1}{2}\left(t - \frac{a_1(a_1+3a_2+2k)}{4t}\right) + \left(\frac{2k+a_1+a_2}{4t}\right)\left(\frac{k}{2} + \frac{a_1-a_2}{4}\right) \\ &= \frac{4k^2+8t^2-a_1^2-6a_1a_2-a_2^2}{16t} \qquad (2.24)\end{aligned}$$

It is easy to show that Nash profits fall in the network benefits of the single-homing side, i.e. $\frac{\partial \pi_{MH}^N}{\partial a_2} = -(3a_1+a_2)/8t < 0$ given that $a_1 > a_2$. Turning to the impact of the multi-homing agents' network benefits a_1, I find that $\frac{\partial \pi_{MH}^N}{\partial a_1} = -(a_1+3a_2)/8t$. Thus, profits fall in a_1 if network benefits are positive on both market sides. In case that the single-homing agents do not benefit from the presence of side-1 agents, however, the negative externality must be small enough in size for Nash profits to fall in a_1.[46] Finally, one can show that Nash profits in case of multi-homing are

[46]Formally speaking, $\frac{\partial \pi_{MH}^N}{\partial a_1} < 0$ if $|a_2| < a_1/3$.

always larger than those under two-sided single-homing given that assumptions 2.1 and 2.3 hold true, because platforms do not directly compete for multi-homing agents.

Let me now turn to the maximum collusive profits that platforms can earn if they collude on both sides of the market. In this case, they will set $p_2^C = k + a_2 n_1^C - t/2$ and choose p_1^C such that it maximizes joint profits on both sides given that demand on side 1 is elastic. In fact, it turns out that platforms already extract the monopoly price from side-1 consumers under Nash competition. Thus, prices on the multi-homing side do not change when allowing for coordination. This result is not surprising in so far as Armstrong and Wright (2007) have already noted that platforms hold a monopoly power over allowing interaction with their single-homing customers (they are so-called "bottlenecks").[47] Summing up, collusive prices and profits are given as follows:

$$p_1^C = \frac{k}{2} + \frac{a_1 - a_2}{4} \; ; \; p_2^C = k - \frac{t}{2} + a_2 \frac{(2k + a_1 + a_2)}{4t} \tag{2.25}$$

$$\pi_{MH}^C = \left(\frac{2k + a_1 + a_2}{4t} \right) p_1^C + \frac{1}{2} p_2^C = \frac{(2k + a_1 + a_2)^2 + 8kt - 4t^2}{16t} \tag{2.26}$$

It is clear that collusive profits must be larger than Nash profits given that nothing changes on the multi-homing side while prices on the single-homing side are increased such that the indifferent consumer located at 1/2 receives a utility of zero. Hence, platforms always have an incentive to collude. Furthermore, this incentive is increasing in the single-homing side's network benefits a_2. For the collusive gain to be a rising function of a_1, however, it must be the case that a_2 is large enough. If network benefits on the single-homing side are actually negative and larger than $(k + a_1)/2$ in absolute value, the gain from colluding falls as a_1 increases.

Before summarizing the above findings in a lemma, let me briefly mention that in case of symmetric externalities $a_1 = a_2 = a$, both Nash and collusive profits are always larger than they were in case of single-homing on both market sides (see

[47] In consequence, platforms would fully extract the utility of multi-homing consumers if transport costs were zero.

section 2.3). Furthermore, π_{MH}^N is falling in a while π_{MH}^C increases in a. Thus, gains from colluding are getting larger as consumers enjoy higher network benefits.

Lemma 2.7 *Suppose that $a_1 > a_2$ and consumers on side 1 are allowed to multi-home. Then, platforms always have an incentive to collude, although profits on the multi-homing side 1 do not increase. Further, $\frac{\partial(\pi_{MH}^C - \pi_{MH}^N)}{\partial a_2} > 0$, but $\frac{\partial(\pi_{MH}^C - \pi_{MH}^N)}{\partial a_1} > 0$ only if $a_2 \geq 0$ or if $a_2 < 0$ and $|a_2| < (k + a_1)/2$.*

A couple of things are worth mentioning with respect to this lemma. First, it turns out that in case of multi-homing, there is no difference between collusion only on one side and collusion on both market sides, a result which is in contrast to section 2.3. This is due to the fact that prices on the multi-homing side are already raised to the monopoly level under Nash. Second, total welfare does not change when moving from Nash to monopoly profits although total demand on the multi-homing side is elastic. This is due to the fact that multi-homing consumers have to pay monopoly prices anyway, thus their demand stays constant when moving from competition to coordinated practices. Yet, consumers on the single-homing side suffer from paying higher prices under collusion implying a decrease in consumer surplus.[48] Finally, I have implicitly assumed that the elastic demand on side 1 is less than one up to now. Given that the mass of consumers on both sides is equal to one, however, side-1 demand is bounded from above also under multi-homing:

$$n_1^N = n_1^C = \begin{cases} \frac{2k+a_1+a_2}{4t} & \text{if } p_1^N > k + \frac{a_1}{2} - t \\ 1 & \text{if } p_1^N \leq k + \frac{a_1}{2} - t \end{cases}$$

In accordance, Nash and collusive profits would have to be adjusted such that side-1 profits simply amount to $1p_1^N$ for all $k \geq 2t - (a_1 + a_2)/2$. I will not analyze this case in more detail, however, because optimal deviation cannot induce a feedback loop in this situation. Instead, the maximum attainable number of multi-homing consumers would have already joined under collusive prices and there would be no

[48] In general, it is therefore not possible to make any statements concerning total welfare implications of coordinated pricing behavior within Armstrong's (2006) framework.

additional gain from defection by attracting more multi-homing consumers. Therefore, I will from now on make the following assumption:[49]

Assumption 2.4 $k < 2t - \frac{a_1+a_2}{2}$

In order to investigate platforms' incentives to deviate from the collusive agreement, I will proceed similarly to section 2.3. More precisely, set $a_1 = a+\Delta$ and $a_2 = a-\Delta$, such that a platform's one-time gain from optimal defection can be analyzed as a function of total network benefits a and an increase in the mean preserving asymmetry Δ. Then, optimal defection prices as functions of collusive prices p_1^C and p_2^C are given as follows:

$$p_1^D = \begin{cases} \frac{(t^2-a^2+\Delta^2)(2k+a+\Delta)-(a+\Delta)(a-\Delta)p_1^C+t(a+\Delta)p_2^C-2atp_2^D}{2(2t^2-a^2+\Delta^2)} & \text{if } k < k_1^* \\ k+(a+\Delta)n_2^D - t & \text{if } k \geq k_1^* \end{cases}$$

$$p_2^D = \begin{cases} \frac{p_2^C+t}{2} + \frac{(a-\Delta)(p_1^C-p_1^D)}{2t} - \frac{a+\Delta}{t}\left(\frac{a-\Delta+p_1^D}{2}\right) & \text{if } k < k_2^* \\ p_2^C - t + \frac{(a+\Delta)(a-\Delta)}{t} + \frac{(a-\Delta)}{t}(p_1^C - p_1^D) & \text{if } k \geq k_2^* \end{cases}$$

$$\text{with } k_1^* = \frac{(8t-a)(t^2-a^2)+(4t-a)\Delta^2}{2t(t+a)+2(t^2-a^2+\Delta^2)} \; ; \; k_2^* = \frac{7(t^2-a^2)+3\Delta^2}{2(t+a)} \; ; \; k_1^* < k_2^*$$

Note that given these deviation prices, deviation demand on the multi-homing side becomes equal to one first, i.e. $k_1^* < k_2^*$. In this case, platforms will set p_1^D on the multi-homing side just high enough to fully extract the utility of a multi-homing consumer located at the opposite end of the Hotelling line, i.e. $p_1^D = k+(a+\Delta)n_2^D - t$. The side-2 deviation price will also adjust since the network benefits of the single-homing consumers do no longer increase because $n_1^D = 1$.[50] Finally, if $k \geq k_2^{**} \equiv \frac{7t}{2} - 2a - \frac{3(a-\Delta)(a+\Delta)}{2t}$, then demand on both market sides will be equal to one. In that case, deviation prices will be given as $p_1^D = k+(a+\Delta)-t$

[49]In case that the assumption is not fulfilled, there will be a range of parameters for which the defector does not gain from lowering prices. Hence, it might be that $\hat{\delta} = 0$. The details on Nash, collusive and deviation profits as well as the critical discount factor in case that $k > 2t-(a_1+a_2)/2$ are found in appendix A.5.

[50]Precisely, the deviation price on side 2 will be given as $p_2^D = \frac{p_2^C+t}{2} - \Delta - \frac{(a-\Delta)(k+a+\Delta)}{2t} + \frac{(a-\Delta)p_1^C}{2t}$ if demand on side 1 is equal to one while demand on side 2 is still smaller than one.

and $p_2^D = k + (a - \Delta) - t - (t^2 - a^2 + \Delta^2)/2t$. The following profit results from the optimal defection strategy just outlined above:

$$\pi_{MH}^D = \begin{cases} \frac{4k(t^2-a^2+\Delta^2)(k+t+a)+8k^2t(t+a)+(t^2-a^2+\Delta^2)^2}{16t(2t^2-2a^2+\Delta^2)} & \text{if } k < k_1^* \\ \frac{2t(t^2-a^2+\Delta^2)(10k+4a-15t)+4t^2(2k(k+2t+2a)-(k-2\Delta)(k+2\Delta))-(t^4-(a^2-\Delta^2)^2)}{16t(2t^2-a^2+\Delta^2)} & \\ & \text{if } k_1^* \leq k < k_2^{**} \\ 2(k+a-t) - \frac{t^2-a^2+\Delta^2}{2t} & \text{if } k_2^{**} \leq k \end{cases}$$

In order to investigate the impact of growing network benefits on platforms' deviation incentives, let me turn directly to the gain of deviation from the collusive agreement:

$$\pi_{MH}^D - \pi_{MH}^C = \begin{cases} \frac{((2k-3(t-a))(t+a)-\Delta^2)^2}{16t(2(t^2-a^2)+\Delta^2)} & \text{if } k < k_1^* \\ \Big[\frac{(a^2-\Delta^2)(5a^2-\Delta^2+8ak+4k^2)-4(a^2-\Delta^2)(2a+3k)t}{16t(2t^2-a^2+\Delta^2)} & \\ -\frac{2(5\Delta^2+2k^2-9a^2)t^2-4(2a+5k)t^3+23t^4}{16t(2t^2-a^2+\Delta^2)}\Big] & \text{if } k_1^* \leq k < k_2^{**} \\ \frac{(4t-k)(k-t+a)+(t-a)(k-3t+a)-2(t^2-a^2+\Delta^2)}{4t} & \text{if } k_2^{**} \leq k \end{cases}$$

Given that assumptions 2.1 to 2.3 hold true, $a, \Delta > 0$ and taking into account the boundaries k_1^* and k_2^{**}, tedious but straightforward calculations show that $(\pi_{MH}^D - \pi_{MH}^C)$ is increasing in the total network benefit a. Hence, a result identical to the one presented in the single-homing setup of section 2.3 can be stated. Turning to the effect of the mean preserving asymmetry Δ, however, I find that the gain from defection actually decreases in Δ, because most of the deviation profits are to be made on the single-homing market side.[51] This second finding is best understood when focusing on the last part of the above equation. In this case, the incentive to deviate is purely driven by price effects as $n_1^D = n_2^D = 1$. Starting with collusive prices, one can see from equation (2.25) that the collusive price on the multi-homing side is an increasing function of Δ. In other words, the larger the difference between network benefits that multi-homing and single-homing consumers enjoy, the higher the additional mark-up to be earned on side 1. Collusive prices on side 2, however, simply depend on the size of the network effect a_2. Thus, the more side-2 consumers

[51]The derivatives are not depicted for the sake of shortness. The corresponding Mathematica files, however, are available from the author upon request.

enjoy the presence of the multi-homing agents, the more they are willing to pay. In turn, this means that p_2^C falls as Δ increases. Turning to the optimal defection strategy, it is easy to note that p_1^D does neither depend on any collusive prices nor on side-2 defection prices. Instead, it simply increases in a_1 given that all single-homing consumers join. Thus, it is an increasing function of Δ. Optimal price defection on side 2, however, depends both on p_1^C and p_2^C as well as on the willingness-to-pay for the network benefit. Overall, the deviation price on side 2 falls by more than p_1^D increases when Δ goes up. Thus, lower network benefits on the single-homing side 1 make deviation less attractive even if the network benefits that the multi-homing consumers enjoy increase by the same amount. Summing up, I can therefore state the following lemma:

Lemma 2.8 *Suppose that $a_1 = a+\Delta$ and $a_2 = a-\Delta$, assumptions 2.3 and 2.4 are fulfilled and consumers on side 1 are allowed to multi-home. Then,* $\frac{\partial(\pi^D_{MH}-\pi^C_{MH})}{\partial a} > 0$ *and* $\frac{\partial(\pi^D_{MH}-\pi^C_{MH})}{\partial \Delta} < 0$.

Wrapping up, the two countervailing effects of increasing network effects a that were found in case of single-homing on both market sides are also present when agents allowed to multi-home. Basically, exactly the same reasoning as outline in section 2.3.1.1 applies. Therefore, I expect the gain from deviation to outweigh collusive gains also in case of multi-homing. Calculating the critical discount factor as given in equation (2.6), I get the following:

$$\hat{\delta}_{MH} = \begin{cases} \frac{(t+a)(2k-3t+3a)-\Delta^2}{(t+a)(2k-5t+5a)+3\Delta^2} & \text{if } k < k_1^* \\ \frac{(a^2-\Delta^2)(4t(2a+3k)-(5a^2-\Delta^2+8ak+4k^2))+2t^2(2k^2-9a^2+5\Delta^2)-4t^3(2a+5k)+23t^4}{(a^2-\Delta^2)(7a^2-3\Delta^2-4k^2+4t(2a+5k))+2t^2(2k^2-27a^2+15\Delta^2-8ak)-4(2a+9k)t^3+47t^4} & \\ & k_1^* \leq k < k_2^{**} \\ \frac{a^2-2\Delta^2-2a(k-4t)-(k-3t)^2}{4a^2-3\Delta^2-(k-6t)(k-2t)+8at} & \text{if } k_2^{**} \leq k \end{cases}$$

Taking first derivatives of $\hat{\delta}_{MH}$, one can show that the critical discount factor increases in a while it falls in Δ. Hence, I can conclude that the incentive to deviate always outweighs the incentive to collude as the total network benefit a increases. For a given total benefit $(a_1 + a_2)$, however, collusion gets easier the more asymmetric network effects on both sides become. Figure 2.11 illustrates the behavior of $\hat{\delta}_{MH}$ with respect to a for different values of Δ.[52]

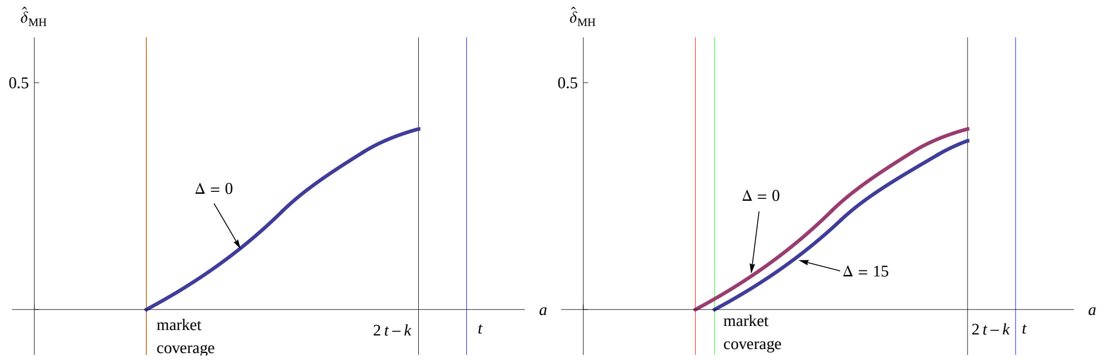
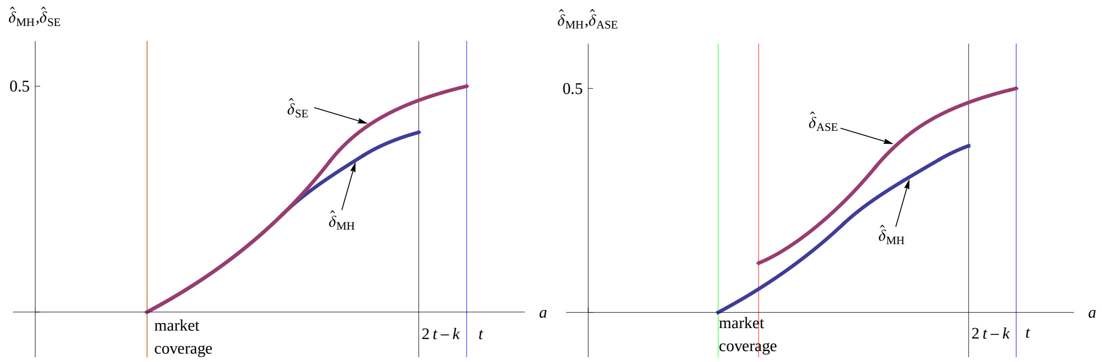


Figure 2.11: critical discount factor in case of multi-homing on side 1

Two things are worth noticing with respect to these graphs: First, $\hat{\delta}_{MH}$ is equal to zero only if assumption 2.3 is binding.[53] Second, the critical discount in case of multi-homing is strictly smaller than 1/2 for all valid parameter ranges. Finally, turning to figure 2.12, it is easy to see that two-sided collusion is always easier to sustain under multi-homing than under two-sided single-homing - a finding which is not surprising given that platforms already charge monopoly prices on the multi-homing side 1 under Nash.



$\Delta = 0$ $\Delta = 15$

Figure 2.12: critical discount factors under single- and multi-homing

I can therefore conclude this section with the following proposition:

Proposition 2.5 *Suppose that $a_1 = a + \Delta$ and $a_2 = a - \Delta$, assumptions 2.3*

[52]Note that both figure 2.11 and 2.12 are drawn for $k = 40$ and $t = 36$.
[53]This corresponds to the vertical green line.

and 2.4 are fulfilled and consumers on side 1 are allowed to multi-home. Then, $\frac{\partial \hat{\delta}_{MH}}{\partial a} \geq 0$ *and* $\frac{\partial \hat{\delta}_{MH}}{\partial \Delta} \leq 0$. *Moreover,* $\hat{\delta}_{MH} \leq \hat{\delta}_{ASE}$.

2.5 Conclusion

This paper is a first step in understanding the impact of indirect network externalities on the sustainability of collusion in two-sided markets. It shows that collusion is harder to sustain when network externalities between the two market sides increase. This is the result of two countervailing effects. First, if one side values members on the other market side more highly, Nash prices fall because competition for this side gets harsher. As a consequence, punishment profits are a falling function of increasing network effects. In addition, consumers' utility from platform participation increases if they enjoy a larger benefit from the presence of platform members on the opposing market side. Hence, two-sided platforms can earn larger collusive profits as network externalities grow. Second and countervailing, however, platforms also earn larger profits from deviation as network effects become stronger - a result which is due to the fact that indirect network externalities induce a feedback loop which leads to more sensitive demand reactions in response to a price decrease. Comparing those opposing effects, it can be shown that the latter effect always dominates in Armstrong's (2006) model of two-sided competition. The reason being that collusion does not change market shares of the colluding parties. Instead, supra-competitive profits are the simple consequence of higher prices. Optimal defection, however, leads to asymmetric platforms sizes on both market sides. Thus, a platform benefits more strongly from the existence of strong network effects under defection than it does when sticking to the collusive agreement.

In addition to this main finding, I show that collusion becomes less attractive if network effects are asymmetric rather than symmetric given that the maximum external benefit that a platform can possibly extract from consumers on both sides stays constant if both consumer groups single-home. Deviation profits rise in response to growing asymmetries in network externalities while the gain from colluding stays constant. When those consumers with higher network benefits are able to multi-home, however, this finding is reversed.

The analysis of partial collusion only on one market side shows that Evans and Schmalensee (2008) have at most been partially right. When firms have to collude on the market side where consumers enjoy higher network effects, they compete away some, but not all, of the supra-competitive profits by setting prices below the competitive level on the other side. If, however, platforms are able to form a cartel on the side of the market that suffers from negative externalities, they might actually make larger additional profits than under cartelization of both market sides. Finally, one-sided collusion might be harder or easier to sustain than collusion on both sides of the market, depending on which side platforms collude on and how big the overall network benefit a is. The general result that collusion becomes harder as the total network benefit increases still holds true.

Coming back to the motivating examples, my results confirm that profitable collusive behavior can take various forms. Importantly, competition authorities should recognize that collusion only on one side can actually be more profitable and easier to sustain than full collusion if one market side imposes negative cross-group externalities. In addition, media markets where only one side has to pay access fees are especially prone to collusion. Finally, evidence for multi-homing indicates that price coordination will be easier given that platforms already charge monopoly-like prices on one side anyway.

Although my analysis of collusion in Armstrong's (2006) often-cited model of competition between two-sided platforms provides interesting insights, it has some caveats. Most importantly, it limits the scope of analysis by not allowing for platforms to charge transaction-based prices. In addition, its focus is on homogeneous membership externalities. Although this form of indirect network effects is well suited to media markets, other two-sided industries might be better described by other forms of network effects. Finally, the chosen framework has to face the general critique addressing Hotelling specifications. In particular, it does not allow for a normative analysis that focuses on total welfare instead of consumer surplus. In fact, even in the multi-homing setup with elastic demand, total surplus stays constant as the platforms already charge monopoly prices from multi-homing consumers in the first place. Thus, future research should focus on generalizations of the demand setup to provide further robustness checks for its main results and derive general welfare implications.

Chapter 3

Complementary Inputs and the Incentives for Upstream Firms to Patent and Innovate

3.1 Introduction

Innovation and imitation go together like a horse and a carriage. It is a well studied and understood phenomenon that technological advances by one firm that push the technological boundaries outward garner substantial interest from competitors. According to Mansfield (1985), it took only about a year for firms to thoroughly understand a new product or process developed by a competitor in the mid-eighties and therefore be able to replicate it. In the meantime, the advances in information acquisition and processing are likely to have sped up the process substantially. Understanding a process or product in combination with the protective boundaries set out in the patent document may actually allow competitors to replicate it without infringing the terms of the patent, so-called "inventing around" existing patents. This has widely been seen as problematic for investment incentives: For example, Gallini (1992) studies the incentives of competitors to invent around existing patents depending on the duration of patent protection, implicitly assuming that this behavior should be preempted.

With the recent shift from a static to a more dynamic depiction of the innovation and patenting process, for example in the context of patent races or cumulative innovations, economists have started to point out potential benefits of a laxer patent policy that does not quench imitation (and inventing around patents) completely. Two studies are particularly important for us in this context: Fershtman and Markovich (2010) study a setting with two firms involved in an R&D-race which has different stages. The firms are asymmetric in the sense that their costs for research differ over the various research stages. In a sense, one could consider the research technologies to be complementary, in that one firm may be better at early research, while the other is better at later research. They numerically uncover equilibria in which the innovation process is overall faster and total welfare is increased if intermediate stages cannot be protected, even if voluntary licensing is possible. Bessen and Maskin (2009) similarly look at innovation processes that can be considered sequential and complementary: Firms can improve an existing product and their innovation costs are independently drawn each period, where they can be either high or low. Complementarity of innovation processes in this context stems from the fact that the costs of the firms are not perfectly correlated; therefore, each period one firm may have high and the other may have low costs of innovation. They find that in this dynamic setting, weakening patent protection can lead to increased total welfare: Intuitively, the originally unsuccessful firm will still have an incentive to invest in the future because this has beneficial consequences, which generally cannot be achieved through licensing because the licensor does not internalize the full effects. Moreover, if past innovations make future progress more likely, even the first innovator (the potential blocking patentee) can benefit from abolishing patent protection.

Our central result is similar to these findings: Ironclad intellectual property right protection can lead to lower investments and lower surplus in a dynamic setting with investments by multiple firms. Yet, the mechanism that we consider is very different to the ones proposed in the two papers above. We show that in a two-period, vertical setting with complementary inputs (instead of complementary innovation technologies, as in the previous papers) ironclad patents can lead to a complete breakdown of upstream investment incentives due to hold-up problems. In the setting with the most distinct results – an ex-ante fixed bargaining order after investments in each period – *every* firm benefits if first-period patents are no longer enforced in the second period.[1] While in the two papers cited above it is the originally less lucky firm whose investment incentives wither away with ironclad patents, in our case it is the upstream firm whose *de facto* bargaining position is weaker, despite the fact that in our setting it is always the investing firms that make take-it-or-leave-it offers, i.e. have full bargaining power. Letting patent protection on earlier investments lapse (which can be interpreted as the possibility to invent around the original patents) improves the bargaining position of the non-residual claimant in the second period. Anticipating this, he also has an incentive to invest earlier on.

We generate these results in the classic contract-theoretical framework of Grossman and Hart (1986); Hart and Moore (1990), adding two (substantial) twists to the setting: First, we consider two upstream suppliers who each carry out an investment into the respective component that they produce. These two components are then combined by the downstream firm D, who we refer to as an assembler, to form the final product, generating the surplus v. The two components are perfect complements in the classical sense: A positive surplus can only be obtained by applying both of them. It appears natural that the total surplus is increasing in each of the investment levels, while one might be able to substitute investments to achieve the same value. Take automobile technology as an example: An interesting exterior design may make up for shortcomings with respect to fuel efficiency (arguably the case for the Chrysler PT Cruiser) in the eyes of customers; or when considering fuel efficiency, the same target can be reached through the newest tire-generation or better engine-technology.

The second twist that we introduce, in the spirit of non-static innovation incentives, is a second period in which the upstream firms again carry out investments and then bargain with D on how the second-period surplus is to be divided. As opposed to the first period, though, now D can choose whether to use the first-period version of

[1] Similarly to Bessen and Maskin (2009), some patent protection is required to ensure investment. In our case, patent protection on investments covers at least the period in which these investments are carried out.

the respective component or the new version for the second-period final product.[2] To stay within the automobile example: During a facelift of a model (or when introducing a new model), the manufacturer can decide for each component whether he wants to employ the latest technology or stick with what was used in the old model. In this context, we can also be more clear about the meaning of ironclad patents vs. knowledge diffusion. Under ironclad patents, the brake producer has the final say both over whether first-generation and whether second-generation brakes are built into the new model. Under knowledge diffusion, a competitive fringe of firms has "invented around" the patents on the first generation brakes, which gives the downstream firm an endogenous outside option.

This second twist of endogenously arising competition for a technology leader over time is perhaps most clearly exemplified in the pharmaceutical industry: Pharmaceutical companies invest to develop prescription drugs, which they aim to protect through patents, in order to be sold via pharmacies after being prescribed by physicians, if necessary. Once a patent expires, generic medication enters the market, usually at a substantially reduced price. To continue to obtain profits, the pharmaceutical companies therefore continuously have to develop and improve products. The trade-off involved in this context is that better "early" products pose a more serious competitive threat to later inventions, which could hamper the follow-up drug's profitability. A first-generation drug which causes no side-effects leaves little room for improvement along this dimension.

What our model enables us to explore is the interaction of incentives arising from (i) the complementarity of inputs and the resulting hold-up problem, and (ii) endogenously arising future competition, which clearly affects the original investment incentives. We believe that this setting is applicable to many different industries. The automobile industry as discussed, but also the computer (with Intel and Microsoft as technology leaders, and large downstream assemblers such as Dell), aerospace, and biotechnology industries. A common feature of all these examples is the central importance of innovation activity at the upstream level, despite the fact that standard theory tells us that both the existence of strong downstream firms and the development of competing new product generations degrade initial investment

[2]In this sense, there exist two overlapping product generations available to D in the second period.

incentives.[3]

Our central results concern the investment incentives of upstream firms in such industries under different intellectual property right protection regimes: We show that investment breaks down completely under ironclad patents. As opposed to this, under imperfect patent protection with knowledge diffusion, both upstream firms invest a strictly positive amount in both periods. Hence, all firms are better off with weaker patent protection. Next, we consider different remedies for these distorted incentives. With respect to the structure of the bargaining game, we replace the predetermined bargaining sequence with a stochastic bargaining schedule. Stochastic bargaining alleviates the holdup-problem in the permanent IP setting and leads to positive investment levels in each period. Incentives to invest in period 2, however, are larger under knowledge diffusion than they are under ironclad patents. Moreover, even period 1 incentives might be larger under knowledge diffusion if the investment incentives of the residual claimant are not distorted too much. Then, we show that partial vertical integration does not resolve the problem of zero innovation in case of ironclad patents. Yet, in case of knowledge diffusion, vertical integration of the downstream firm with one of its suppliers results in higher investment levels than non-integration. Finally, we show that the only situation leading to first-best investment levels is when both upstream suppliers merge and first-period innovations are protected by long-lasting IP rights, a solution which resembles the classical result of Grossman and Hart (1986).

The remainder of this paper is organized as follows. The next section shortly discusses the related literature. Section 3.3 describes our theoretical model and derives first-best investment levels. Then, we investigate investment incentives under non-integration when patents are ironclad as well as when they expire after period 1 (see section 3.4). As a robustness check of our findings, we allow for different bargaining procedures. Next, section 3.5 studies the impact of partial vertical integration as well as horizontal upstream integration on upstream innovation incentives. Section 3.6 compares investment levels and profits under all ownership scenarios and derives

[3]The latter argument is known since Coase (1972) and has more recently been studied in the context of innovation incentives by Fishman and Rob (2000) or Nahm (2004), for example. For a comprehensive review of the first argument, namely the effects of buyer power, see e.g. Inderst and Mazzarotto (2008).

policy implications. Section 3.7 concludes.

3.2 Related Literature

In addition to the research described in the introduction, our paper is related to two different strands of literature. First, it is in line with standard contract theory in the sense of Grossman and Hart (1986); Hart and Moore (1990). In their seminal article, Grossman and Hart (1986) analyze a static model with two players, each making one relationship-specific investment and owning control over their respective asset. Ex-ante, only ownership is contractible, but assets are contractible ex-post and renegotiation is allowed. In contrast to our setup, Grossman and Hart (1986) assume that there exists a competitive market of potential trading partners at date zero, which determines the ex-ante division of the surplus. Given this generally distortive setup, they show that player 1 should own both assets if his ex-ante investment is much more important than firm 2's, i.e. in case that firm 2's under-investment is relatively unimportant and over-investment by player 1 is less severe than under-investment by player 1 in case of non-integration would be. In contrast, non-integration, i.e. the situation where each agent owns his asset, will be desirable if both investments are important because this leads to intermediate investments undertaken by both agents. Based on these results, Hart and Moore (1990) enrich the setup to treat an arbitrary number of agents and assets. Each of the agents makes an investment before production and trade occur. Assuming cooperative bargaining over the surplus realization, gains will be split according to the Shapley value. To determine an agent's marginal return on investment, Hart and Moore (1990) assume that investments are complementary at the margin and that the marginal return on investment increases both with the number of other agents and assets in the coalition that an agent belongs to. They show that strictly complementary assets should always be owned together – a result which is reproduced in our model only if IP rights are ironclad. Moreover, they demonstrate that partial integration in a setup with 3 agents and assets might increase or decrease the incentives to invest of the stand-alone agent depending on how the surplus depends on individual investments. For example, agent 3 will have a greater investment incentive under integration of 1 and 2 if either 1 and 3 obtain limited synergy in the absence of 2 or if 2 and 3 obtain limited synergy in absence of 1.[4]

Given that the above seminal articles boosted research in contract theory, we limit our attention to those papers that focus on applications to innovation incentives within a (vertical) industry setup. Aghion and Tirole (1994) consider a setup with two players, one research unit and one customer, that both undertake relationship-specific investments into one asset (the innovation). They show that the Grossman-Hart result is still valid, i.e. there will be under-investment by both parties. Moreover, they show that the allocation of property rights is always efficient when the research unit has ex-ante bargaining power, whereas the fact that the research unit is cash constrained might lead to an inefficient ownership structure when the customer has ex-ante bargaining power. Moreover, there are several papers investigating two sequential investments and showing that simple option contracts solve the hold-up problem in these settings if renegotiation is allowed before the second investment takes place (see Nöldeke and Schmidt (1998), Lülfesmann (2001)). Smirnov and Wait (2004a) show that sequencing can allow some projects to proceed that would not have been feasible with simultaneous investments. Yet, sequencing might disadvantage some parties leading to lower investments. In consequence, the possibility of sequencing might reduce welfare or even discourage investment of the first-mover completely, thus preventing trade from occurring. In line with this finding, Smirnov and Wait (2004b) show that the timing of investment can act as an additional form of hold-up if investing parties choose a welfare-reducing time regime. Our main result is related to these findings in the sense that investment levels are distorted in case of overlapping product generations and bargaining after each period.

Finally, our paper also relates to the recent industrial organization literature on the impact of vertical integration (VI) and buyer power on upstream innovations. Inderst and Wey (2003) investigate different mergers in an industry setup with two innovating upstream suppliers and two downstream retailers operating in different regional markets. Assuming that industry profits are split up between all indepen-

[4]Note, however, that the bargaining setup in Hart and Moore (1990) avoids any over-investment by assumption. Thus, there will be some under-investment for any ownership structure (see their proposition 1).

dent suppliers and retailers according to a multilateral bargaining procedure that results in Shapley values, they derive conditions under which up- and downstream firms choose to merge. In case of linear cost and demand functions, upstream incentives to adopt a new low-cost technology are higher if upstream firms are separated while downstream firms merge. Thus, an increase in buyer power increases overall welfare. In our paper, we show that a higher outside option of the downstream firm in the second period leads to positive upstream investments in case of knowledge diffusion. When buyer power stays constant at zero because of ironclad patents, however, there will be no investments at all. In contrast to Inderst and Wey (2003), who focused on upstream process innovation, Ishii (2004) allows for both upstream and downstream innovation. In particular, he analyzes the effects of cooperative R&D in vertically related duopolies with knowledge spillovers. He shows that technological improvement is accelerated by vertical research joint ventures. Thus, social welfare is largest if vertically related firms can coordinate their R&D decisions and/or fully share useful knowledge. This is in line with our finding that partial VI of the downstream firm with one of its suppliers increases investment levels under knowledge diffusion as compared to those under non-integration. Buehler and Schmutzler (2008) examine endogenous VI and cost-reducing downstream R&D in a linear Cournot model. They find that VI has an intimidation effect, i.e. it increases own investment and decreases competitor investment. Hence, complete vertical separation is less likely in presence of R&D investments than in a benchmark without investments. This is in line with our finding that partial VI increases innovation incentives and profits of both the remaining independent supplier and the merged entity, given that the "solo" supplier still has full bargaining power when selling his investment.

Turning to the relation between buyer power and innovation incentives, the traditional view has been that increased buyer concentration reduces upstream investment incentives. Inderst and Wey, however, have qualified these findings in a series of papers inspecting both process and product innovation (see Inderst and Wey (2005, 2007, 2011)). In Inderst and Wey (2005), they show that powerful buyers might induce a supplier to invest more into cost reductions because this increases his fraction of incremental profits and at the same time reduces the attractiveness of a large buyer's outside options. In a follow-up paper, Inderst and Wey (2007) show that if a supplier faces large buyers, his bargaining position depends crucially on how well he can deal with the loss of market share after disagreement leading to a strong incentive to flatten his cost function or increase his capacity via investments. In fact, such an investment might lead to more output and increased welfare. What our paper has in common with these findings is once again that increased downstream buyer power due to knowledge diffusion in period 2 leads to higher investment levels than an outside option of zero in case of ironclad patents, as long as investments are carried out by two distinct upstream parties. Finally, we would like to mention that an advantage of our approach as compared to the IO papers mentioned above is that our general results hold true for both process and production innovations.

3.3 The Model

Structure and Timing We consider a simple two-period game with overlapping product generations. For a depiction of the timeline, see **figure 3.1**. In each period, a single downstream firm D sells a product which is composed of two components (or requires two inputs), each of which is produced by an upstream firm. In period one, the two upstream firms, each of whom produces one of the components $j \in \{1, 2\}$, simultaneously carry out investments i_1, i_2 at costs $c(i)$. For simplicity, we assume that the marginal costs for each type of investment in both periods are constant and equal to $c > 0$. These investments can be interpreted, e.g., as the innovative effort spent on designing the component that the supplier produces. If the two components are combined, this results in the period one total value (or revenues) $v(i_1, i_2)$, which is twice differentiable, strictly concave and strictly increasing in i_j. On its own, each input is useless and has a value of 0. Thus, the downstream firm requires both components and has an outside option of 0. Their period one innovations grant the suppliers full bargaining power vis-à-vis D: A potential interpretation is that they are able to fully protect their investments through patenting and therefore own the property rights of their inventions in the sense of Grossman and Hart (1986).

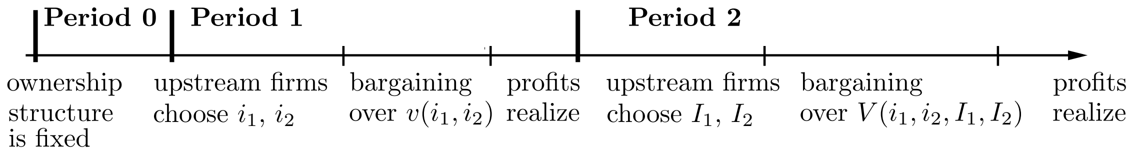


Figure 3.1: timing of the game

In the second period, the proceedings from the first period are basically repeated, but with an important difference. The upstream firms again invest into a new generation of components, simultaneously choosing investment levels I_1, I_2.[5] If both new components are employed, this results in the period two total value $V(i_1, i_2, I_1, I_2)$, which is twice differentiable, strictly concave and strictly increasing in each of the arguments i_j, I_j. Further, we assume that a modified Inada-condition on the limit of the first partial derivatives with respect to both inputs holds true:[6]

Assumption: $\lim_{i_i, I_i \to 0} V'(.) = \infty$ *if* $i_j > 0$ *or* $I_j > 0$ *with* $i \neq j;\ i, j \in \{1, 2\}$

Given that the two inputs are complements, it is natural to assume that the marginal product of the first unit of investment into component i is very large as long as a positive level for any of the two product generations of the other input j exists.

The fundamental difference to the previous period is that now two generations of each component exist (if both i and I are strictly positive): To produce a product in the second period, D can therefore choose between the new and the old input

[5] Throughout, we use lowercase variables for the first, and uppercase variables for the second period.

[6] Note that we do not make any assumptions concerning the cross-derivatives of $V(.)$. Coming back to our automotive example given in the introduction, a plausible assumption would be that $\frac{\partial I_i}{\partial I_j} < 0$, $\frac{\partial I_i}{\partial i_j} < 0$, but $\frac{\partial I_i}{\partial i_i} > 0$, with $i \neq j; i, j \in \{1, 2\}$. Another interesting, yet very restrictive, assumption is $\frac{\partial I_i}{\partial i_j} = 0$. In this case, the second period investment of supplier 1 would not depend on the prior investment of the complementary supplier 2. If this latter assumption is fulfilled, the first-order conditions on investment incentives to be derived in sections 3.4 and 3.5 would simplify substantially.

of both products. One of our central interests is how intellectual property rights affect the incentives of the firms to invest in innovation. For this, we consider two regimes:

1. On the one hand, a setting which we refer to as **technology** or **knowledge diffusion**: Imagine that between the periods a competitive fringe of suppliers has managed to reproduce the period one inventions, while circumventing the existing patents. As a result, the downstream firm can appropriate the rents resulting from the use of the old technology. In other words, its period two outside option is endogenously given by $V(i_1, i_2, 0, 0)$.

2. On the other hand, a setting in which the patents protecting the period one investments are ironclad, which we refer to as **permanent intellectual property (IP) rights**. As a result, each firm remains the only one to offer components of its type and the outside option of the downstream firm stays equal to zero in period two.

In either IP-setting, after their investment, the upstream firms bargain with D regarding the application of period one or period two technology, with the upstream firms making "take-it-or-leave-it" offers. Finally, we solve the above described two-stage game for subgame-perfect Nash equilibria by backward induction.

We present the model in detail as follows: First we analyze the case in which all three firms (upstream firms 1 and 2 and the downstream firm D) are independent entities, which we refer to as the assembler-case, because all that firm D does is buy the two components and assemble them to a final product, without any relevant investment decision on its part. Contract-theoretically, this setting resembles the case of a buyer of two products that are perfect complements, whose quality depends on the investments of their producers. This case is interesting in itself, as we will see that first-best cannot be achieved in the (simple) sequential three party bargaining setting which we impose, despite the fact that the investing players have full bargaining power in the sense of take-it-or-leave-it offers. As a first important result, we show that investment even breaks down completely under the permanent IP rights regime, while both upstream firms invest a strictly positive amount in both periods under technology diffusion. Next, we consider different remedies for

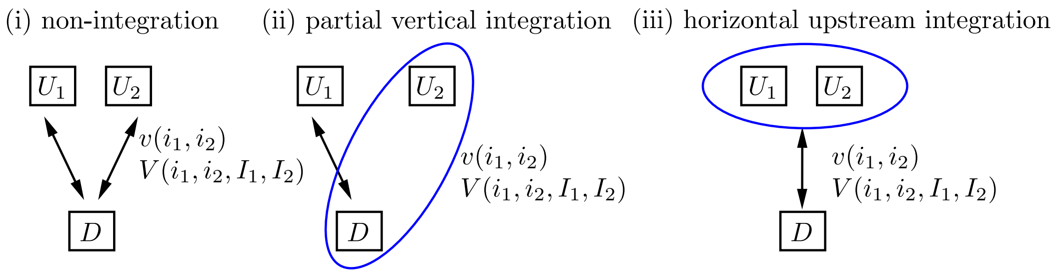


Figure 3.2: possible industry structures

the distorted incentives encountered in the assembler case. The first is with respect to the structure of the bargaining game, replacing the predetermined bargaining sequence with a stochastic bargaining schedule.[7] In addition to the robustness of the bargaining game, we also analyze two different ownership settings, as depicted in **figure 3.2**. We show that under partial upward integration of D, which we call the manufacturer case, investment still breaks down under permanent IP rights, while under knowledge diffusion, investment levels increase over the ones in the assembler case. Finally we show that first-best investment can only be achieved under horizontal upstream integration with permanent IP rights, which resembles the classical result of Grossman and Hart (1986). As a benchmark for our analysis, however, let us start by stating first-best investment levels of the above described game.

First-best investment levels A social planner maximizes total surplus with respect to first and second period investments:

$$\max_{i_1,i_2,I_1,I_2} \quad v(i_1,i_2) - c(i_1) - c(i_2) + \beta[V(i_1,i_2,I_1(i_1,i_2),I_2(i_1,i_2)) - c(I_1(i_1,i_2)) - c(I_2(i_1,i_2))] \tag{3.1}$$

First-best investments in period 2 for a given investment choice in period 1 are therefore given by the levels I_1^{FB} and I_2^{FB} that satisfy the following first-order

[7]This can be seen as the more general case, for which the non-stochastic order is a subcase. We proceed as described both for expositional simplicity and because one may think of cases in which it is hard to find an interpretation for the stochastic setting.

conditions, respectively:

$$\frac{\partial V(i_1, i_2, I_1, I_2)}{\partial I_1} = c \tag{3.2}$$

$$\frac{\partial V(i_1, i_2, I_1, I_2)}{\partial I_2} = c \tag{3.3}$$

Turning to first-period investments given $I_1^{FB}(i_1, i_2)$ and $I_2^{FB}(i_1, i_2)$, we can apply the envelope theorem to obtain the following first-order conditions that implicitly define i_1^{FB} and i_2^{FB}:

$$\frac{\partial v(i_1, i_2)}{\partial i_1} + \beta \left[\frac{\partial V(i_1, i_2, I_1^{FB}(i_1, i_2), I_2^{FB}(i_1, i_2))}{\partial i_1} \right] = c \tag{3.4}$$

$$\frac{\partial v(i_1, i_2)}{\partial i_2} + \beta \left[\frac{\partial V(i_1, i_2, I_1^{FB}(i_1, i_2), I_2^{FB}(i_1, i_2))}{\partial i_2} \right] = c \tag{3.5}$$

3.4 The Assembler Case

In order to investigate the setup with two suppliers and one downstream firm, we first need to fix a non-cooperative bargaining procedure for the three players that are involved. The upstream firms make sequential take-it-or-leave-it offers to D. We first present the regime in which the order of offers is fixed, with firm 1 making the first offer in each period, before generalizing to a stochastic order of offers, in which firm i correctly perceives the probability of itself making the first offer in period t as $p_i^t \in (0, 1)$ at the time of the investment decision.

3.4.1 Fixed Order Sequential Bargaining

Technology Diffusion We first turn to the IP-rights setting in which there is diffusion of knowledge, so that the upstream firms do not have sole ownership of their period 1 investments in period 2. In each period, upstream firms 1 and 2 make simultaneous investments. After this, first firm 1 makes a take-it-or-leave-it offer to D, then firm 2 makes such an offer. When bargaining takes place in period 2, all investments i_1, i_2, I_1, I_2 are fixed. Therefore, the total period 2 revenues to be distributed amount to $V(i_1, i_2, I_1, I_2)$, with D able to obtain $V(i_1, i_2, 0, 0)$ on its own.

The lemma below provides the equilibrium outcome for the period 2 negotiations (see appendix B.1 for proof).

Lemma 1: *The following is the unique subgame-perfect Nash equilibrium of sequential bargaining in period 2 under knowledge diffusion: Firm 1 demands* $B_1 = V(i_1, i_2, I_j, 0) - V(i_1, i_2, 0, 0)$*, firm 2 demands* $B_2 = V(i_1, i_2, I_1, I_2) - V(i_1, i_2, I_j, 0)$ *with* $j = 1$ *and* D *accepts both offers.*

There is an intuitive interpretation for this result: In the first negotiation, success leads to the formation of a coalition of supplier 1 and D. This coalition generates the value $V(i_1, i_2, I_1, 0)$ on its own, while failure of negotiations would lead to D realizing $V(i_1, i_2, 0, 0)$. Therefore the upstream firm making the take-it-or-leave-it offer can (and does) request $V(i_1, i_2, I_1, 0) - V(i_1, i_2, 0, 0)$. Similarly, the second negotiation with firm 2 decides whether the coalition producing the good in the second period is composed of $\{D, 1\}$, or $\{D, 1, 2\}$. If the negotiation is successful, the surplus $V(i_1, i_2, I_1, I_2)$ is generated. Without 2, on the other hand, the surplus is $V(i_1, i_2, I_1, 0)$. The difference between the two is the sum that 2 can request in its take-it-or-leave-it offer.

Next, we turn to bargaining in the first period. Again, assume that after simultaneous investments firm 1 makes the first take-it-or-leave it offer to D, followed by firm 2. It is worthwhile to note that investments i_1, i_2 are sunk at this point in time and that $V(.)$ depends on the investment levels in the first period, but not on the bargaining outcome.[8] The following lemma then provides the equilibrium outcome for the period 1 negotiations (see appendix B.2 for proof):

Lemma 2: *The following is the unique subgame-perfect Nash equilibrium of sequential bargaining in period 1 under knowledge diffusion: Firm 1 demands* $b_1 = 0$*, firm 2 demands* $b_2 = v(i_1, i_2)$ *and* D *accepts both offers.*

We get this relatively extreme result due to the structure of the bargaining game with the two firms demanding fixed-sum payments from D and because of the assumption that both inputs are required to generate a positive value – in a sense,

[8]Subgame-perfection allows us to abstract from punishment strategies contingent on bargaining outcomes.

firm 2 is able to exploit this and hold-up both D and firm 1.[9] Given these bargaining outcomes, we are able to spell out the profit functions of the three firms under non-integration and sequential bargaining:

$$\begin{aligned}\pi_1 &= -c(i_1) + \beta[V(i_1, i_2, I_1, 0) - V(i_1, i_2, 0, 0) - c(I_1)] &(3.6)\\ \pi_2 &= v(i_1, i_2) - c(i_2) + \beta[V(i_1, i_2, I_1, I_2) - V(i_1, i_2, I_1, 0) - c(I_2)] &(3.7)\\ \pi_D &= \beta V(i_1, i_2, 0, 0) &(3.8)\end{aligned}$$

Turning to the investment decisions, we first focus on whether an equilibrium can arise in which no investments whatsoever are carried out. Firm 1's period 2 investment problem immediately reveals that $I_1 = 0$ must be chosen whenever $i_2 = 0$, while a positive level is chosen if $\frac{\partial V(i_1,i_2,0,0)}{\partial I_1} > c$, which is assured by the Inada-type assumption regarding $V(.)$. Analogously, firm 2 invests a positive amount in the second period whenever $i_1 > 0$ or $I_1 > 0$. Now consider firm 2's investment incentives in the first period, starting with the worst case, i.e., the reaction to $i_1 = 0$, with all incentives to invest being exerted by potential earnings in period 2: firm 2 anticipates that a positive investment level in period 1 will give firm 1 a reason to invest in the second period, which in turn allows it to earn a positive profit then.[10] Therefore, the Inada-conditions on V allow us to focus on interior solutions for which the first-order conditions hold.

We derive upstream investments by backward induction in two steps. First, we solve for profit maximizing investment levels in the second period given i_1, i_2:

$$\frac{\partial V(i_1, i_2, I_1, 0)}{\partial I_1} = c \qquad (3.9)$$

$$\frac{\partial V(i_1, i_2, I_1, I_2)}{\partial I_2} = c \qquad (3.10)$$

Given period one investments i, it is immediately apparent that firm 2 has socially

[9] The following subsection which allows for stochastic bargaining orders can be seen as a natural robustness analysis.

[10] The same reasoning holds for i_1, in turn.

efficient investment incentives, maximizing V in its decision.[11] This is associated with the insight from Lemma 2 that firm 2, making the last offer in the sequential bargaining game, can be considered the residual claimant given the three-player bargaining structure that we impose. With respect to I_2, the payment that firm 1 obtains is a lump-sum payment, that does not distort the investment incentives of firm 2. On the other hand, the investment incentives of firm 1 are distorted downwards compared to the social optimum. Intuitively, firm 1 only internalizes the gains to the smaller coalition including itself and D.

Given these implicitly defined optimal investment levels $I_1^*(i_1, i_2)$ and $I_2^*(i_1, i_2)$, we turn to the profit-maximizing first period investments, given by the following first-order conditions on i_1, i_2:

$$\text{FOC}_1^I: \qquad \beta\left[\frac{\partial V(i_1,i_2,I_1^*(i_1,i_2),0)}{\partial i_1} - \frac{\partial V(i_1,i_2,0,0)}{\partial i_1}\right] = c \tag{3.11}$$

$$\begin{aligned}\text{FOC}_2^I: \qquad & \frac{\partial v(i_1,i_2)}{\partial i_2} + \beta\left[\frac{\partial V(i_1,i_2,I_1^*(i_1,i_2),I_2^*(i_1,i_2))}{\partial i_2} - \frac{\partial V(i_1,i_2,I_1^*(i_1,i_2),0)}{\partial i_2}\right. \\ & \left. + \left(\frac{\partial V(i_1,i_2,I_1^*(i_1,i_2),I_2^*(i_1,i_2))}{\partial I_1} - \frac{\partial V(i_1,i_2,I_1^*(i_1,i_2),0)}{\partial I_1}\right)\frac{\partial I_1}{\partial i_2}\right] = c \end{aligned} \tag{3.12}$$

It is straightforward to see that firm 1's investment will be lower than the social optimum. Since it does not receive any surplus in the first period, it only internalizes the increase in the share of period 2 surplus which it can claim on its own. On the other hand, since firm 2 can be seen as a kind of residual claimant in both periods, its period 1 investment incentives are relatively close to the optimum. It obtains the entire period 1 surplus, and with regard to the second period it internalizes firm 1's reaction to its higher investment.

Permanent IP rights Now consider the analogous setting under the permanent IP rights regime. Now, each firm has sole control over the utilization of the component that it produces, no matter which period. This simplifies the game substantially: Lemma 2 is unchanged as the patenting regimes in the first period are identical. Only the second period bargaining game is affected. By a parallel

[11] See Section 3.3 for the social optimum.

argument to above, firm 2 will demand the entire surplus $V(i_1, i_2, I_1, I_2)$, while firm 1 cannot achieve a positive payout with D accepting the offer. Backward induction shows that investment incentives break down completely in this setting: Neither firm invests a positive amount in either period.[12] We consolidate these results in the proposition below:

Proposition 1: *Consider the assembler case with two upstream firms investing in consecutive periods into components that resemble perfect complements for a final product offered by a downstream firm. Given the assumptions on V(.), in the fixed order sequential bargaining setting no positive investment can be achieved under permanent intellectual property rights, while under knowledge diffusion strictly positive investments are carried out by both firms in each period, which constitutes a Pareto-improvement.*

Note that this result can easily be generalized to n upstream firms: Under knowledge diffusion, each firm would receive the marginal benefit it generates for the respective coalition it joins in the second period, while first period investments (except for the firm that is the last in the bargaining sequence) are solely motivated by the prospect of benefiting in the second period.

3.4.2 Stochastic Order Sequential Bargaining

While in the previous section we studied a case of bargaining in which the order of offers is fixed, we now turn to a setting in which there is uncertainty regarding the order of bargaining at the time that investments are made. While under the fixed order, given our assumptions, the last firm to bargain can be considered critical and is able to hold-up both other parties, in a stochastic setting it is unclear ex-ante

[12] Note that this result hinges on the fact that firm 1 is the first to bargain in both periods. If bargaining order would reverse in the second period, both firms would choose positive investment levels in period 1. In particular, if firm 1 goes first in period 1 and firm 2 bargains first in period 2, the following investment levels will be chosen: I_1 will be such that $\frac{\partial V(i_1,i_2,I_1,0)}{\partial I_1} = c$, $I_2 = 0$, i_1 satisfies $\beta \left[\frac{\partial V(i_1,i_2,I_1,0)}{\partial i_1}\right] = c$ and i_2 satisfies $\frac{\partial v(i_1,i_2)}{\partial i_2} = c$. Note, however, that all investment levels except for i_1 will be unambiguously higher under knowledge diffusion than under permanent IP even with this alternating bargaining procedure. Moreover, also i_1 will be higher under knowledge diffusion if:
$$\frac{\partial V(i_1,i_2,I_1,I_2)}{\partial i_1} - \frac{\partial V(i_1,i_2,0,I_2)}{\partial i_1} + \left(\frac{\partial V(i_1,i_2,I_1,I_2)}{\partial I_2} - \frac{\partial V(i_1,i_2,0,I_2)}{\partial I_2}\right)\frac{\partial I_2}{\partial i_1} > \frac{\partial V(i_1,i_2,I_1,0)}{\partial i_1}$$

who will be the one to wield the most influence. One can interpret the fixed-order case as one with a strong hierarchy among suppliers, while the stochastic approach would apply to a more symmetric setting.

Recall that we defined the probability of firm i making the first offer in a given period, perceived at the time of investment, i.e. ex ante, as $p_i^t \in (0,1)$. Therefore $p_2^t = 1 - p_1^t$. Further, take the offers b and B from Lemmas 1 and 2, which resemble the offers that a firm makes when making its offer first or second, respectively.[13] Again, we begin with the knowledge diffusion IP right setting. Then, the firms' expected profits are given as follows:

$$\pi_1 = (1-p_1^1)b_2 - c(i_1) + \beta[p_1^2 B_1 + (1-p_1^2)B_2 - c(I_1)] \tag{3.13}$$

$$\pi_2 = (p_1^1)b_2 - c(i_2) + \beta[p_1^2 B_2 + (1-p_1^2)B_1 - c(I_2)] \tag{3.14}$$

$$\pi_D = \beta V(i_1, i_2, 0, 0) \tag{3.15}$$

From these profit functions, we derive optimal investments analogously to the previous section. We obtain I_1^* and I_2^* as the investment levels that solve the following first-order conditions:

$$p_1^2 \frac{\partial V(i_1,i_2,I_1,0)}{\partial I_1} + (1-p_1^2)\frac{\partial V(i_1,i_2,I_1,I_2)}{\partial I_1} = c \tag{3.16}$$

$$p_1^2 \frac{\partial V(i_1,i_2,I_1,I_2)}{\partial I_2} + (1-p_1^2)\frac{\partial V(i_1,i_2,0,I_2)}{\partial I_2} = c \tag{3.17}$$

Based on $I_1^*(i_1,i_2)$ and $I_2^*(i_1,i_2)$, we can derive the first-order conditions that implicitly characterize i_1^* and i_2^*:

$$\begin{aligned}(1-p_1^1)\tfrac{\partial v(i_1,i_2)}{\partial i_1} + \beta \Big[p_1^2 \Big(\tfrac{\partial V(i_1,i_2,I_1^*(i_1,i_2),0)}{\partial i_1} - \tfrac{\partial V(i_1,i_2,0,0)}{\partial i_1}\Big) \\ +(1-p_1^2)\Big(\tfrac{\partial V(i_1,i_2,I_1^*(i_1,i_2),I_2^*(i_1,i_2))}{\partial i_1} + \tfrac{\partial V(i_1,i_2,I_1^*(i_1,i_2),I_2^*(i_1,i_2))}{\partial I_2}\tfrac{\partial I_2}{\partial i_1} \\ - \tfrac{\partial V(i_1,i_2,0,I_2^*(i_1,i_2))}{\partial i_1} - \tfrac{\partial V(i_1,i_2,0,I_2^*(i_1,i_2))}{\partial I_2}\tfrac{\partial I_2}{\partial i_1}\Big)\Big] = c\end{aligned} \tag{3.18}$$

[13]Note that this notation is a little sloppy for the sake of short notations. It is crucial to understand that $j = 1$ in lemma 1 only if firm 1 makes the first offer, while $j = 2$ if firm 2 makes the first offer.

$$p_1^1 \frac{\partial v(i_1,i_2)}{\partial i_2} + \beta \Big[(1-p_1^2)\Big(\frac{\partial V(i_1,i_2,0,I_2^*(i_1,i_2))}{\partial i_2} - \frac{\partial V(i_1,i_2,0,0)}{\partial i_2}\Big)$$
$$+p_1^2\Big(\frac{\partial V(i_1,i_2,I_1^*(i_1,i_2),I_2^*(i_1,i_2))}{\partial i_2} + \frac{\partial V(i_1,i_2,I_1^*(i_1,i_2),I_2^*(i_1,i_2))}{\partial I_1}\frac{\partial I_1}{\partial i_2}$$
$$-\frac{\partial V(i_1,i_2,I_1^*(i_1,i_2),0)}{\partial i_2} - \frac{\partial V(i_1,i_2,I_1^*(i_1,i_2),0)}{\partial I_1}\frac{\partial I_1}{\partial i_2}\Big)\Big] = c \tag{3.19}$$

It is straightforward to see that each of these conditions on i_1^* and i_2^* is a convex combination of equations (3.11) and (3.12). The probabilities p shift the likelihood that a firm will be the one to be able to hold up the others, which increases (or decreases) the incentives to invest in either period.

Given the stochastic problem, it is a natural extension to consider how the levels of p are determined in each period. Imagine that D, who is buying the two components in order to assemble them, can act as an agenda-setter and choose the order of bargaining in both periods. Also note that apart from being able to decline any given offer, this is the only choice that D can actively make in this setting. Therefore, we are able to define D's problem as:

$$\max_p \beta V(i_1^*, i_2^*, 0, 0) \tag{3.20}$$

As D's period 2 outside option is the only source of rents to the firm, it will attempt to maximize the value created by period 1 investments of the upstream firms, while disregarding the effects on period 2 investments. In general, this will induce additional distortions, so that we can straightforwardly state the following proposition:

Proposition 2: *Consider the assembler case with stochastic bargaining and knowledge diffusion. If D can set the bargaining agenda, second-best is not achieved unless:*

$$\arg\max_p \beta V(i_1^*, i_2^*, 0, 0) = \arg\max_{p'} v(i_1^*, i_2^*) + \beta V(i_1^*, i_2^*, I_1^*, I_2^*) \tag{3.21}$$

Next, we consider the alternative IP rights setting with permanent patents. We derive the expected profit functions of the firms as well as the resulting first-order conditions in appendix B.3. From these, it is immediately clear that, all else given, incentives for period 2 investments are higher under knowledge diffusion. The intuition for this is that firms have efficient investment incentives if they are the residual claimant in either setting, while their incentives are strictly higher under knowledge diffusion in the case that the opposite firm is the residual claimant. Both firms' investment incentives increase when the hold-up potential via period 1 patents is diffused in the second period.

What is more surprising are the trade-offs involved between the settings regarding period 1 investment incentives. One might have expected that strict protection unanimously increases period 1 investments given an interior solution. In fact, the intuition just discussed partially carries over. If the firm is not the residual claimant in the second period (i.e., if the firm has to make the first offer), then the fact that it can still make some profit in the knowledge diffusion case will also increase its incentive to invest in the first period. What is now different, though, is that investments are distorted downwards under knowledge diffusion in the case that the firm is the residual claimant: In their first period decisions, the upstream firms anticipate that the bargaining position of the firm moving first is affected by their period 1 decision, which cuts into its own profits. We can derive the differences between the first-order conditions for the two firms in case of knowledge diffusion versus ironclad patents. Given supermodularity, i.e. if $\frac{\partial^2 V(.)}{\partial i_j \partial I_j} > 0$ for $j \in \{1,2\}$, both period 1 *and* period 2 investment incentives are stronger under knowledge diffusion if the following conditions are fulfilled:

$$\Delta\text{FOC}_1: \quad p_1^2 \left(\frac{\partial V(i_1,i_2,I_1^*(i_1,i_2),0)}{\partial i_1} - \frac{\partial V(i_1,i_2,0,0)}{\partial i_1} \right) \\ -(1-p_1^2) \left(\frac{\partial V(i_1,i_2,0,I_2^*(i_1,i_2))}{\partial i_1} + \frac{\partial V(i_1,i_2,0,I_2^*(i_1,i_2))}{\partial I_2} \frac{\partial I_2}{\partial i_1} \right) > 0 \tag{3.22}$$

$$\Delta\text{FOC}_2: \quad (1-p_1^2) \left(\frac{\partial V(i_1,i_2,0,I_2^*(i_1,i_2))}{\partial i_2} - \frac{\partial V(i_1,i_2,0,0)}{\partial i_2} \right) \\ -(p_1^2) \left(\frac{\partial V(i_1,i_2,I_1^*(i_1,i_2),0)}{\partial i_2} + \frac{\partial V(i_1,i_2,I_1^*(i_1,i_2),0)}{\partial I_1} \frac{\partial I_1}{\partial i_2} \right) > 0 \tag{3.23}$$

We can therefore summarize the above findings in the following proposition:

Proposition 3: *Consider the assembler case with stochastic bargaining. Positive investment levels will be achieved under both IP rights regimes. Given the assump-*

tions on $V(.)$ and assuming supermodularity of first and second period investments, investment incentives will be higher under knowledge diffusion than under permanent IP rights if conditions (3.22) and (3.23) are fulfilled.

Finally, it is important to point out the source of profits for the downstream firm as an assembler. It can only achieve a positive rent under the knowledge diffusion setting, as only then its period 2 outside option is greater than zero. The thought experiment that leads to Proposition 2 clearly demonstrates the incentives of the otherwise passive assembler: He tries to push for a large amount of early innovation in order to profit from it later on. Our results would also predict an inclination of assemblers to try to undermine the intellectual property rights of suppliers to induce knowledge diffusion. Coming back to the automotive industry example given in the introduction, a famous case in point would be the spread of the ESP-technology developed by Bosch, which was accelerated by German downstream manufacturers.[14]

3.5 Alternative Ownership Settings

To complete our analysis, we compare the assembler case with overlapping product generations to two other ownership settings. This demonstrates the trade-offs in detail and relates it to more well-known contractual settings. In particular, we analyze the situation in which D has merged with one of the suppliers and the situation in which the two upstream suppliers have merged horizontally.

[14]Note that the German automobile industry is a good example of our assembler setup given that more than 75% of total value creation and the larger share of innovation activities is carried out by suppliers (see Verband der Automobilindustrie (2011)). Bosch is the largest automobile supplier worldwide based on revenues, with a particular knowledge in brake systems. In 2010, the Bosch Group invested more than 3.8 billion euros into R&D and applied for more than 3,800 patents worldwide (see `http://www.bosch.de/start/content/language2/html/867.htm`). Other complementary and also highly innovative upstream suppliers of German automotive manufacturers are, for example, ZF Friedrichshafen AG (driveline and chassis technology), Mahle Group (engine systems and filtration) and Continental (tires, interior and safety).

3.5.1 Vertical Integration: The Manufacturer Case

First, we analyze the case in which D merges with one of the upstream firms, calling the remaining upstream entity "firm 1". In this case, we are in a more standard two-player bargaining setting. As before, we assume that the upstream firm, as the owner of a critical patent, has full bargaining power in the sense that it makes a take-it-or-leave-it offer. Note that proposition 1 carries over immediately, i.e. the merged entity will have no incentive to invest under permanent IP rights. Therefore, we only study the knowledge diffusion case.

First, consider the period 2 bargaining stage where all investment levels are fixed. Given that the vertically integrated firm can realize profit $V(i_1, i_D, 0, I_D)$ on its own, the stand-alone upstream firm 1 will make a take-it-or-leave-it offer demanding $B_1 = V(i_1, i_D, I_1, I_D) - V(i_1, i_D, 0, I_D)$ and D accepts. In period 1, firm 1 will make D indifferent between accepting and declining its offer in the first period by demanding the entire period 1 surplus through $b_1 = v(i_1, i_D)$ and D accepts. Therefore we obtain the following profit functions for the two firms:

$$\pi_1 = v(i_1, i_D) - c(i_1) + \beta[V(i_1, i_D, I_1, I_D) - V(i_1, i_D, 0, I_D) - c(I_1)] \quad (3.24)$$

$$\pi_D = -c(i_D) + \beta[V(i_1, i_D, 0, I_D) - c(I_D)] \quad (3.25)$$

The solution is analogous to the procedure in section 3.4. We start by solving for profit maximizing second period innovation levels I_1^* and I_D^* given i_1 and i_D:

$$\frac{\partial V(i_1, i_D, I_1, I_D)}{\partial I_1} = c \quad (3.26)$$

$$\frac{\partial V(i_1, i_D, 0, I_D)}{\partial I_D} = c \quad (3.27)$$

We find that D's second period investment incentives are identical to those of the firm making the first offer in section 3.4.1. The upstream firm, as the residual claimant, has identical period 2 incentives to the firm making the last offer in 3.4.1. The period-2 distortions arising in the manufacturer case are therefore identical to those in the fixed order sequential assembler case.

Next, we turn to the first period decisions and derive the first-order conditions that implicitly characterize i_1^* and i_D^*:

$$\text{FOC}_1^{II}: \quad \frac{\partial v(i_1,i_D)}{\partial i_1} + \beta \left[\frac{\partial V(i_1,i_D,I_1^*(i_1,i_D),I_D^*(i_1,i_D))}{\partial i_1} - \frac{\partial V(i_1,i_D,0,I_D^*(i_1,i_D))}{\partial i_1} + \left(\frac{\partial V(i_1,i_D,I_1^*(i_1,i_D),I_D^*(i_1,i_D))}{\partial I_D} - \frac{\partial V(i_1,i_D,0,I_D^*(i_1,i_D))}{\partial I_D} \right) \frac{\partial I_D}{\partial i_1} \right] = c \quad (3.28)$$

$$\text{FOC}_D^{II}: \quad \beta \left[\frac{\partial V(i_1,i_D,0,I_D^*(i_1,i_D))}{\partial i_D} \right] = c \quad (3.29)$$

Again, the upstream firm has identical incentives to the assembler case. What changes is the investment incentive of the merged firm as opposed to the separate upstream firm: Since D now internalizes the improved outside option in the second period completely, it has a stronger incentive to invest in the first period than in the assembler case. A vertical merger therefore increases the investment incentives of the upstream firm being acquired, while it does not directly affect the incentives of the remaining firm, if this firm has full bargaining power.[15]

Thus, we can state the following proposition:

Proposition 4: *Consider the manufacturer case with the stand-alone firm 1 having full bargaining power. Then, permanent IP rights lead to zero investments in both periods. Under knowledge diffusion, the vertically merged entity has higher investment incentives than under non-integration, while investment levels of the stand-alone firm stay constant.*

3.5.2 Horizontal Upstream Integration: The Distributor Case

Finally, we turn to the case in which both upstream firms are integrated. In this scenario, all investment decisions are made by the single supplier U. In both periods, U makes a take-it-or-leave-it offer to D after choosing innovation levels 1 and 2. Given that the downstream firm cannot make any profit without both invest-

[15] Clearly, we hereby disregard potential increases in bargaining power of the merged entity, which might be expected in reality.

ments in period one, its outside option is zero and U will demand $b_U = v(i_1, i_2)$. In the second period, however, the downstream firm has a positive outside option $V(i_1, i_2, 0, 0)$, given that patents on the old innovations have expired in case of knowledge diffusion. Thus, U can only demand the additional surplus generated with the second period innovations: $B_U = V(i_1, i_2, I_1, I_2) - V(i_1, i_2, 0, 0)$.[16] Given these bargaining outcomes, profit functions are as follows:

$$\begin{aligned} \pi_U &= v(i_1, i_2) - c(i_1) - c(i_2) \\ &\quad +\beta[V(i_1, i_2, I_1, I_2) - V(i_1, i_2, 0, 0) - c(I_1) - c(I_2)] \end{aligned} \tag{3.30}$$

$$\pi_D = \beta V(i_1, i_2, 0, 0) \tag{3.31}$$

In order to derive investment incentives of the horizontally integrated upstream supplier, we proceed once again in two steps. It turns out that for given first period investments i_1, i_2, second period investments will be efficient, that is $I_1^*(i_1, i_2)$ and $I_2^*(i_1, i_2)$ are given by the following first-order conditions:

$$\frac{\partial V(i_1, i_2, I_1, I_2)}{\partial I_1} = c \tag{3.32}$$

$$\frac{\partial V(i_1, i_2, I_1, I_2)}{\partial I_2} = c \tag{3.33}$$

First period investment levels maximize upstream profits taking into account that second period investments are ex-post efficient. Hence, we derive the following implicit functions of i_1^* and i_2^*:

$$\text{FOC}_1^{III}: \quad \frac{\partial v(i_1,i_2)}{\partial i_1} + \beta\left[\frac{\partial V(i_1,i_2,I_1^*(i_1,i_2),I_2^*(i_1,i_2))}{\partial i_1} - \frac{\partial V(i_1,i_2,0,0)}{\partial i_1}\right] = c \tag{3.34}$$

$$\text{FOC}_2^{III}: \quad \frac{\partial v(i_1,i_2)}{\partial i_2} + \beta\left[\frac{\partial V(i_1,i_2,I_1^*(i_1,i_2),I_2^*(i_1,i_2))}{\partial i_2} - \frac{\partial V(i_1,i_2,0,0)}{\partial i_2}\right] = c \tag{3.35}$$

It is straightforward to see that period one investment incentives are distorted. In fact, U internalizes the improved bargaining position of D in period 2 that would result from higher investment levels i_1 and i_2.[17]

[16]Note that the assumption that $V(i_1, i_2, I_1, I_2)$ strictly increases in each of the arguments guarantees that U's profit is maximized for positive levels of both second period innovations, i.e. $V(i_1, i_2, I_1, I_2) > V(i_1, i_2, I_i, 0)$ with $i \in \{1, 2\}$.

[17]This result is in fact very close to the durable goods monopolist problem, only extended to

Comparing the above first-order conditions with those of section 3.4.1, we find that second period investment levels are higher than in the assembler case. Moreover, the period-1 investment incentive of the upstream firm that was first to bargain in the assembler case increases after a horizontal upstream merger. Turning to the supplier that is the residual claimant under sequential bargaining in 3.4.1, its incentives to invest will only be higher under horizontal integration if the following holds true.

$$\text{FOC}_2^{III}\text{-FOC}_2^{I}: \quad \left(\frac{\partial V(i_1,i_2,I_1^*(i_1,i_2),0)}{\partial i_2} - \frac{\partial V(i_1,i_2,0,0)}{\partial i_2}\right) - \left(\frac{\partial V(i_1,i_2,I_1^*(i_1,i_2),I_2^*(i_1,i_2))}{\partial I_1} - \frac{\partial V(i_1,i_2,I_1^*(i_1,i_2),0)}{\partial I_1}\right)\frac{\partial I_1}{\partial i_2} > 0 \tag{3.36}$$

Finally, let us check what happens under horizontal upstream integration when patents are ironclad. In this case, D will have an outside option of zero in both periods, because he cannot generate any positive surplus without buying from the integrated upstream supplier. Thus, investment incentives are not distorted by the endogenous outside option of D. Thus, this setup implements first-best investments as given in section 3.3.

We can therefore state the following:

Proposition 5: *Consider the distributer case. Then, permanent IP rights lead to first-best investments in both periods. Under knowledge diffusion, horizontal upstream integration leads to first-best investment incentives in period 2. Incentives to invest in period 1 also increase above those under non-integration if condition (3.36) is fulfilled.*

3.6 Comparison of Investment Incentives and Profits

Table 3.1 summarizes our findings for the different industry structures.

two "goods" or innovations (see Coase (1972)).

	(i) non-integration (fixed-order bargaining)	(ii) partial vertical integration	(iii) horizontal upstream integration
knowledge diffusion	$0 < i_j^I, I_j^I < i_j^{FB}, I_j^{FB}$	$I_1^I = I_D^{II}, I_2^I = I_1^{II}$; $i_2^I = i_1^{II}$; $i_1^I < i_D^{II}$	$I_j^{III} = I_j^{FB}$; $i_1^I < i_1^{III}$; $i_2^I \lesseqgtr i_2^{III}$
ironclad IP rights	$i_j^I, I_j^I = 0$ with $j = \{1, 2\}$	$i_j^{II}, I_j^{II} = 0$ with $j = \{1, D\}$	$i_j^{III} = i_j^{FB}$; $I_j^{III} = I_j^{FB}$ with $j = \{1, 2\}$

Table 3.1: Investment incentives

It is straightforward to see that first-best investment can only be achieved with ironclad patents and upstream horizontal integration - a finding which is in line with standard contract theory (Grossman and Hart (1986)). When horizontal upstream integration is not possible, however, e.g. because of antitrust concerns, it is Pareto-superior to have knowledge diffusion instead of permanent IP rights in our model. While no investment will be made under ironclad IP rights in settings (i) and (ii), knowledge diffusion leads to positive investments by both firms in each period.

Moreover, D always prefers technology diffusion as this is the only way to earn positive profits $V(i_1, i_2, 0, 0)$. Turning to the question which industry setup yields the highest surplus for D, we find that if condition (3.36) is fulfilled, D will prefer to have one integrated instead of two independent and complementary suppliers. Comparing settings (i) and (ii), upstream firm 1 always chooses a higher first-period investment under horizontal integration than under non-integration, while firm 2 will do so only if the indirect effect that his investment i_2 has on its second period surplus via I_1 is smaller than the direct effect of i_2 on the surplus of its competitor firm 1 in case (i). Finally, it is not clear if D prefers vertical integration over the other two setups without making additional assumptions on profit sharing within the vertically merged entity. Yet, we know that the investment incentives of the vertically merged entity are higher than the incentives of the non-integrated upstream firm 1 in the assembler setup. Thus, it follows right away that joint profits of the vertically merged manufacturer are larger than the sum of profits of D and firm 1 under non-integration.

A final interesting observation with respect to table 3.1 concerns the choice of industry structure. If condition (3.36) holds and (FOC_2^{III}-FOC_2^{II})> 0, then letting firm 1, who is first to bargain under non-integration, decide if and to whom to propose a merger will always lead to the highest possible investment levels both under permanent IP rights and under knowledge diffusion.

3.7 Concluding Remarks

In this paper, we study a contract theoretic setting in which two upstream firms make investments into components that are perfect complements to produce a final product in two periods. As a new twist, in the second period, the final product can be composed of one of each components from either period - that is, a new component of one type can be substituted by the old component of the same type. We show that this introduces distortions to the investment incentives in the spirit of a durable goods monopolist: Suppliers anticipate that they will be competing against their old product in the next period if patents are not ironclad. On the other hand, it is exactly this structure that can contribute to the hold-up problem being alleviated: If an endogenous outside option exists in the second period, then each of the upstream firms can claim a part of the total surplus that corresponds to the value of their contributions to the respective coalition with D. The ability of the residual claimant to exploit the hold-up situation is thereby reduced. We also show that this attribute translates to the more general case of stochastic-order bargaining under certain conditions. Finally, we show how our setting with two complementary investments in two periods relates to more "standard" cases by studying the effects of both vertical and horizontal integration.

Chapter 4

The Compromise Effect in Action: Lessons from a Restaurant's Menu

4.1 Introduction

Standard discrete choice analysis assumes that consumers behave rationally when choosing an option from a given choice set. More precisely, they maximize their utility of consumption by picking the option that is ranked highest given their individual preference structure (see e.g. Luce (1959), McFadden (1973)). Hence, a crucial assumption of rational choice theory is that consumers have a well-defined preference structure over all possible choice alternatives. This implies that a change in choice set composition does not lead to choice reversal between options available in both sets.

However, many laboratory experiments, mostly in psychology and marketing research, but also in experimental economics, have shown that individuals do not necessarily behave according to standard theory when facing changes in the choice set.[1] In particular, a number of "context effects" have been documented repeatedly (see e.g. Simonson (1989), Kahneman, Knetsch, and Thaler (1991), Tversky

[1] For a summary of recent developments in non-standard decision making, see e.g. DellaVigna (2009), Section 4. An excellent review of the history of choice theory, refinements of the basic model, the psychology of choice and applicable statistical methods is given by McFadden in his Nobel lecture (McFadden (2001)).

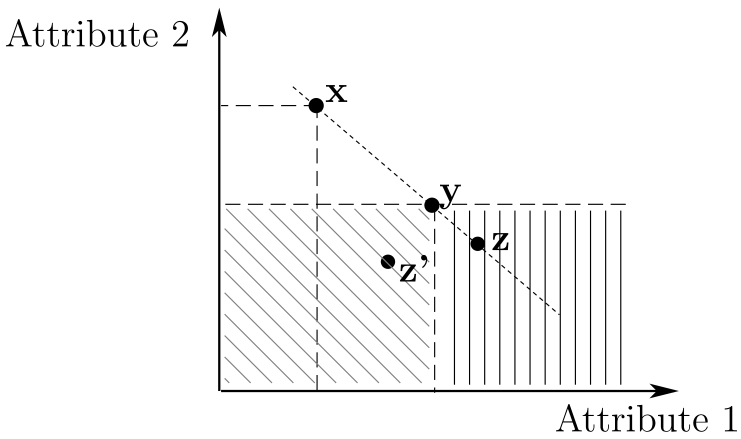


Figure 4.1: Attraction vs. compromise effect

and Simonson (1993) and Sheng, Parker, and Nakamoto (2005)).[2] Context effects imply that individuals' preferences are not fixed but dependent upon the context of choice. In other words, the utility that a decision maker assigns to an option depends not only on its own characteristics, but is the result of considering the properties of all options in the choice set.

One of the most well-known context effects is the so-called "compromise effect", documented first by Simonson (1989). It refers to a situation where a choice alternative gains market share when it moves from being an *extreme option* to a *compromising* or middle option with respect to all relevant choice attributes. In an example with 3 options and 2 attributes as depicted in Figure 4.1, this means that y's choice probability increases if an option in the vertically hatched area (e.g. option z) enlarges the choice set $\{x, y\}$.[3] As a possible explanation for the compromise effect, Simonson (1989) states that "[...] choice behavior under preference

[2]Please note that the theory of context effects has to be distinguished from the literature on pure "framing effects" (see e.g. Salant and Rubinstein (2008); Salant (2011)). In all the above mentioned studies as well as in our data (see Section 4.3), the actual composition of the choice set varies, whereas a change in the "framing" of a choice situation refers to a variation of the ordering or presentation of elements in the choice set without changing the elements and their attributes themselves.

[3]The compromise effect has to be distinguished from the related but not identical observation of an attraction or asymmetric dominance effect. According to Huber, Payne, and Puto (1982), the latter describes the observation that adding a new option which is strictly dominated by one but not all other options already in the choice set leads to an increase in the choice probability of the dominating option. In Figure 4.1, this corresponds to the addition of z' to a set consisting of x and y, where z' has to lie in the diagonally hatched rectangle. As y strictly dominates z', its choice probability increases according to the attraction effect.

uncertainty may be easier to explain by assuming that consumers select the alternative supported by the best reasons", such as the enhancement of self-esteem or the anticipation of possible regret. As a consequence, individuals prefer an option whose attribute levels lie in between those of the remaining choices, because it is the easiest to justify under uncertainty about attribute importance.[4]

Several authors proved the robustness of Simonson's observation by enrichment of the basic experimental setup. Variations in the experimental design included the complexity of the choice set, the familiarity with the product or the influence of time pressure (Lehmann and Pan (1994), Kivetz, Netzer, and Srinivasan (2004a), Sheng, Parker, and Nakamoto (2005), Dhar, Nowlis, and Sherman (2000)). However, all evidence so far has been obtained from laboratory experiments. The question whether compromise effects manifest in purchase decisions in the field is more difficult to answer and has so far not been investigated.[5]

This article contributes to the existing literature by being the first to provide "real world" evidence for the existence of a compromise effect. To this end, we construct a completely new and unique data set from raw data provided by a German specialties restaurant. Overall, we observe more than 88,000 individual choices of main courses from 21 different menus offered over a period of more than 7 years. In these menus, main dishes are grouped into 6 categories such as fish, steaks or vegetarian food, and are listed within each category in an ascending-price order. Hence, we can easily classify choice options as being on the extreme end or serving as a compromise along a price-quality trade-off line.[6] Variation in price, position and number of

[4]de Clippel and Eliaz (2010) propose a nice formalization of the reason-based choice argument made by Simonson (1989). According to their model, a decision maker has several "inner selves" that represent the different attributes of the available options. These conflicting inner selves will reach a compromising choice as the unique solution of a standard bargaining process.

[5]There are, however, two studies checking the existence of an attraction effect in field decisions (see Doyle, O'Connor, Reynolds, and Bottomley (1999) and Josiam and Hobson (1995)). Both studies find mixed evidence. In consequence, they argue that for the attraction effect to work in reality, the seller has to signal convincingly that one option is definitely inferior to another - a requirement that might be hard to fulfill in more complex and uncertain choice situations in natural, non-laboratory, environments. Furthermore, there exists a large literature on extremeness aversion in survey design, i.e. with respect to the use of Likert scales (see e.g. O'Muircheartaigh, Krosnick, and Helic (2000)).

[6]In line with many experimental studies (e.g. Simonson, Nowlis, and Lemon (1993), Dhar, Nowlis, and Sherman (2000), Lehmann and Pan (1994)), we claim that a product's characteristics can be subsumed under the notion of overall quality and price. According to Green and Srinivasan (1978) and others, customers often face market choices with a trade-off between price and several

items within these 6 different dish category choice sets from one menu to another allows us to investigate the effects of several "pure" cases of choice set expansion as well as to estimate various discrete choice models.

Our findings indicate that the compromise effect prevails both in descriptive and regression analyses. It is more pronounced when an alternative switches from being the most expensive item (instead of being the lowest-price item) in the choice set to being a compromise option. Moreover, our results indicate that controlling for confounding factors such as the background context of individual decision makers is important, as it can change the size of the compromise effect, without however changing its sign or statistical significance.

Our results are relevant in at least two respects. First, they are of interest to theoretical researchers, who aim at building empirically accurate models for economic choice behavior. Second, they are important to practitioners and marketers, who may exploit the compromise effect to construct choice sets strategically, such that the attractiveness and purchase likelihood of designated (high-margin) options is maximized. Our empirical investigation of compromise effects in real purchases thus has significant implications for the positioning and branding of products, as well as for competitive strategies in general.[7]

The remainder of this paper is organized as follows. Section 4.2 formally defines the concept of a compromise effect and reviews the most important theoretical explanations for this phenomenon. Section 4.3 shortly presents the new data set we constructed, while Section 4.4 describes our empirical strategy. Our results are discussed in Section 4.5 and Section 4.6 concludes with implications and limitations of this study.

---

quality related dimensions that are highly correlated. Customers are then likely to simplify such choices by making their decision on the basis of price versus overall product quality. While the attribute "price" is well-defined in our data set, we claim that the quality dimension of a main dish is highly correlated with its price. Accordingly, a dish with a lower price must be of lower quality than another one exhibiting a higher price. This seems to be reasonable as price setting, according to the owner of the restaurant, is based on a cost-plus-markup method. Hence, more expensive ingredients as well as more sophisticated preparation - both being indicators of higher quality - determine higher prices.

[7]For anecdotal evidence that this type of choice behavior is already exploited by some companies, see the advertisements posted in Kivetz, Netzer, and Srinivasan (2004a), p. 239.

4.2 Theoretical Framework

This section first defines the notion of a compromise effect within the standard choice framework and derives implications that are testable using aggregate data. Then, it shortly summarizes the theoretical literature on compromise effects.

According to most standard (probabilistic) choice models (see e.g. Luce (1977), McFadden (1973)), a rational decision maker bases his choice on utility maximization over a finite choice set containing all available options. In consequence, observable choices should be congruent with certain axioms of rational choice. Let us define $T = \{x, y, z...\}$ as the finite set of all options and let C be a choice function associating with any offered set $S \subset T$, i.e any nonempty subset of T.[8] Then, $C(S)$ includes all options in S that are chosen by the decision maker. Utility maximization implies that there exists a function v assigning a real value to each x in T such that $x \in C(S)$ if and only if $v(x) \geq v(y)$ for all $y \in S$. This implies that the ordering of options will be independent of the choice set, i.e. if $x \in C(S)$ and $x \in R \subset S$ then $x \in C(R)$. In other words, a non-preferred option cannot become preferred when new options are added to the choice set, a property that is commonly known as "independence of irrelevant alternatives (IIA)" (Luce (1959)).

It is important to note that the axioms of rational choice, such as the above described IIA, cannot always be tested when only aggregate choice measures such as market shares are available. There are, however, two conditions applicable to aggregate data whose violation would indicate that individuals behave irrationally: a stronger one based on absolute choice probabilities and a weaker one based on relative probabilities. The first one is the so-called "regularity condition" (see e.g. Luce (1959), Huber, Payne, and Puto (1982)). Let $P(x; S)$ be the proportion of people for whom $x \in C(S)$ and assume that ties between values of different choice options are excluded.[9] Then, the regularity condition is fulfilled if $x \in R \subset S$ implies that $P(x, R) \geq P(x, S)$. Hence, the market share of an option does not increase after an expansion of the choice set. This can easily be verified using aggregate data. The second and much weaker condition in terms of irrational behavior is based on relative choice probabilities. It was derived by Tversky and Simonson

[8] This and the following notation is borrowed from Tversky and Simonson (1993).

[9] For the ease of notation, we will write $P(x; y)$ for $P(x; \{x, y\})$ and $P(x; y, z)$ for $P(x, \{x, y, z\})$.

(1993) under some additional structural assumptions and is commonly known as "betweenness inequality".[10] Assuming that a choice option is described by a vector of distinct attributes such as price and quality and that preferences are monotonic and separable in attributes, they define an option y to lie between x and z if and only if for every attribute i it holds that $x_i \leq y_i \leq z_i$ or $x_i \geq y_i \geq z_i$. Then, utility maximization implies that the middle option y becomes less popular relative to x if an additional option z is introduced. Formally, Tversky and Simonson (1993) define the market share of y relative to x as

$$P_z(y;x) = \frac{P(y;x,z)}{P(y;x,z) + P(x;y,z)} \tag{4.1}$$

Hence, if $x_i \leq y_i \leq z_i$ or $x_i \geq y_i \geq z_i$ for all i, then utility maximization[11] implies that $P(y;x) \geq P_z(y;x)$.

As already explained in the introduction, several experimental studies have shown that the implications of utility maximization are often violated in actual choice situations. Choice set expansions often lead to unpredicted choice behavior due to the compromise effect (see e.g. Simonson (1989); Gaertner and Xu (1999)). More precisely, the introduction of a new choice option leads to a *violation of the betweenness inequality or even the regularity condition* for those alternatives in the original choice set that become a compromising option after the set expansion. In response to the recurrent documentation of the effect, psychologists have come up with many reasons for compromising behavior, e.g. the avoidance of regret in case of choice uncertainty, prevention and promotion motivations, group acceptance or loss aversion. While most of these arguments are made informally, some studies in marketing research and behavioral economics present theoretical models for choice behavior leading to a violation of the IIA property.

Following his seminal paper in 1989, Itamar Simonson develops a formal model of context-dependent preferences joint with Amos Tversky (see Simonson and Tversky

[10] Note that in terms of a logical test of individual rationality, betweenness inequality would be seen as a stronger criterion than the regularity condition. It is the condition that is more easily violated if individuals do not behave according to rational choice theory.

[11] In addition to the above restrictions, another assumption called ranking condition is needed to prove betweenness inequality. According to most authors, the ranking condition is fulfilled in most applications. See the appendix in Tversky and Simonson (1993) for more details.

(1992); Tversky and Simonson (1993)). They propose that the context of choice influences decision making via two different channels. First, according to the "local trade-off contrast hypothesis", the addition of a new option might increase the market share of the existing option with the highest relative advantage: If people compare each option to all other available alternatives and make their choice based on the aggregate relative advantage of each option, then this trade-off contrast behavior implies a violation of betweenness inequality. Second, Simonson and Tversky (1992) extend the concept of loss aversion such that advantages and disadvantages are no longer defined with respect to a choice-neutral reference point, but in relation to the context of choice, i.e. the available alternatives. This is what the authors call "extremeness aversion". Consider once again the example of a choice set with three options characterized by a two-dimensional attribute space as given in Figure 4.1. If y lies in between the options x and z, it offers only small advantages and disadvantages in relation to the two extreme alternatives. Options x and z, on the other hand, exhibit large advantages on one attribute dimension, while being highly unattractive on the other dimension when compared to one another. In comparison with the middle option y, both extremes have a small advantage on one dimension as well as a small disadvantage on the other. Aggregation of pairwise comparisons under the assumption that relative disadvantages are weighted more highly than the corresponding advantages leads to a preference of the compromising middle option when all three are available. Formally, $P_x(y;z) > P(y;z)$ and $P_z(y;x) > P(y;x)$.[12] Hence, extremeness aversion is inconsistent with utility maximization and the resulting betweenness inequality. Yet, it does not necessarily violate the regularity condition.

Formalizing the above arguments, let $C_B(.)$ be the choice function associated with local context B.[13] Then, $C_B(S) = x$ means that, given B, x is chosen from choice

[12]Simonson and Tversky (1992) further distinguish between compromise and polarization effects. A compromise effect is given when both inequalities hold, i.e. when loss aversion is true with respect to both attributes. In contrast, they speak of polarization when only one dimension is prone to loss aversion and henceforth only one of the two inequalities holds true.

[13]In addition to the local context, Tversky and Simonson (1993) also include the background context into their model indicating that even prior choice situations might influence a decision maker's actual choice from the target set. As we have no information on former individual choice situations in our data set, we abstract from this component of the model (see Tversky and Simonson (1993), p. 1184ff for details.)

set S. Given this notation, an option x will be chosen if its utility as derived from the following context dependent, additive function $V_B(x, S)$ is larger than the utility of any other alternative in the choice set, whereby:[14]

$$V_B(x,S) = \underbrace{v(x)}_{\substack{\text{context-}\\ \text{free}\\ \text{value}}} + \underbrace{\theta g(x,S)}_{\substack{\text{impact of}\\ \text{the choice}\\ \text{set}}} \quad \text{with } \theta \geq 0 \tag{4.2}$$

Tversky and Simonson (1993) expect θ to be positive when choice is difficult or uncertain. In order to derive the compromise effect, some further assumptions are made. First, the binary preference order between two options x and y is additive in all n attributes, i.e. $x \succ y$ if and only if $v(x) = \sum_{i=1}^{n} v_i(x_i) > \sum_{i=1}^{n} v_i(y_i) = v(y)$. Second, the choice set S influences the value of x via pairwise comparisons of the advantages $A(x, y)$ and disadvantages $D(x, y)$ of x with respect to all $y \in S$.[15] Hence, the second component of equation (4.2) will be a function of the relative advantage of x over all other options available whenever there are more than two alternatives in the choice set. Formally, Tversky and Simonson (1993) define:

$$g(x,S) = \begin{cases} \sum\limits_{y \in S} R(x,y) & \text{if } s > 2 \\ 0 & \text{if } s \leq 2 \end{cases} \tag{4.3}$$

$$\text{with } R(x,y) = \frac{A(x,y)}{A(x,y) + D(x,y)}$$

where s denotes the number of alternatives in S. Given this additional structure,

---

[14] Note that this equation reduces to standard utility maximization if $\theta = 0$.

[15] The advantage of x over y for a given attribute i is defined as:

$$A_i(x,y) = \begin{cases} v_i(x_i) - v_i(y_i) & \text{if } v_i(x_i) \geq v_i(y_i) \\ 0 & \text{otherwise.} \end{cases}$$

The overall advantage follows by adding up the attribute-wise advantages, i.e. $A(x, y) = \sum_{i=1}^{n} A_i(x, y)$. The disadvantage $D_i(x, y)$ of x over y is defined as an increasing convex function δ_i of $A_i(y, x)$ with $\delta_i(t) \geq t$.

equation (4.2) can be rewritten as the following "componential context model":

$$V_B(x,S) = \sum_{i=1}^{n} v_i(x_i) + \theta \sum_{y \in S} R(x,y) \tag{4.4}$$

Using the latter component of the model, the authors show that an individual who is indifferent between x and y in a binary choice set might choose the middle option y from $\{x, y, z\}$ if $R(y, z) > R(x, z)$, that is if the relative advantage of y over z is larger than the one of x over z. In the aggregate measure, this "trade-off contrast" leads to a violation of regularity. In a similar manner, one can show that extremeness aversion leads to a preference for the compromise option even if the context-free values of x, y and z are identical.[16]

Another argument explaining the existence of the compromise effect is choice uncertainty. Wernerfelt (1995) claims that most consumers are uncertain about their exact valuation for a certain choice option although they are able to rank their valuation within the population of decision makers. If the available choice set reflects the taste distribution within this population at least to some degree, then a consumer may use observable market data to infer his own valuation for a certain option. Accordingly, he should pick the option that consumers with his taste rank would buy. Following such behavior can lead to a violation of regularity if consumers that first face a small and then a larger choice set are uncertain about the population ranking in the smaller choice set. In contrast, Sheng, Parker, and Nakamoto (2005) hypothesize that consumers have to make a decision under uncertainty about the performance of a product. This asks for a choice process that accounts for the risk of disappointment. Sheng, Parker, and Nakamoto (2005) presume that decision makers pick the option that minimizes the expected loss with respect to a certain reference point. If a consumer is uncertain about his correct reference point, i.e. the alternative that yields the highest utility, it might be optimal for him to choose the compromise option. Last, an alternative approach to explain compromise effects is presented in Bordalo (2010). He shows that if decision makers focus only on salient attributes and these salient attributes depend on the choice set, then compromise effects can arise.

[16]For the detailed derivation of these results, see Tversky and Simonson (1993), p.1186-87.

Finally, there exists one study that compares the performance of different context-dependent choice models including the one by Tversky and Simonson (1993) presented above. Kivetz, Netzer, and Srinivasan (2004a) present four different predictive models. They all describe a constructive decision making process based on a relative and an absolute utility component, thereby allowing for standard utility maximization as a special case of the respective model when the context-dependent relative component vanishes. In two empirical applications, they test which of the four models performs best to predict the observed compromise effect in their choice experiments. To this end, they calibrate the models to match individual-level attribute valuations and weights, which have been elicited from questionnaires handed out to participants before the choice experiment. Results indicate that normalized single-reference point models have superior aggregate predictive validity[17] and yield bigger improvements with respect to the standard utility maximization model than the model of Tversky and Simonson (1993). In addition, these models outperform Tversky and Simonson's (1993) model with respect to model fit as measured by the Schwarz Bayesian information criterion. Hence, the authors conclude that single-reference point models are superior to the tournament-like model proposed by Tversky and Simonson (1993), in which each option in the set serves as a reference point for all other options. A possible drawback of all tested models, however, is that they are of a cardinal nature and assume additivity as well as a particular functional structure for the calculation and aggregation of individual attribute levels (de Clippel and Eliaz (2010)).

Summing up, several authors have proposed explanations for the compromise effect observed in laboratory experiments. Many theoretical models have been developed in which individuals behave according to a new form of constructive, context-dependent preferences and in which at least some uncertainty is involved in the choice process. Concerning the empirical identification of the best-fitting theory, there is only one paper indicating that single-reference point models that account for the local context of choice might be most appropriate to explain the compromise effect. Yet, not all arguments outlined above have been tested against one another. Further evidence is needed about which models provide the best explanation for

[17] This validity is measured by the mean absolute deviation of the observed shares in the validation sample from the predicted choice shares based on the calibration sample.

"irrational" choice behavior observed in reality. Yet, no single model is likely to capture reality to the fullest. Therefore, it is useful to keep all possible explanations in mind when turning to the empirical analysis in the following sections as this might help us to understand the sensitivity of results with respect to the inclusion of different covariates. Or, as Dekel and Lipman (2010) point out: "Intuitively, a given model is like a particular explanation of the behavior it generates. It is unsurprising that at least some choice behaviors may have multiple explanations. Furthermore, multiple explanations may be useful. Different explanations will suggest different intuitions, different questions to consider, different comparatives that might be useful."

4.3 The Data

Our final choice dataset consists of three different sources that were merged after a thorough data cleaning process.[18] In detail, we use:

1. Electronic cashier system data from 01/05/2002 until 05/29/2009
2. Pdf-files of all menus handed out during the above period
3. Official data on local weather conditions, unemployment rates, price indices and data from the CESIfo Group Munich on the "Ifo Business Climate Germany"

The German specialties restaurant, which kindly provided us with the first two data sets, is located in a rural area of North-Eastern Germany.[19] Data source 1 includes all the restaurant's bills. Extractable information comprises alternative-specific data, such as the amount and exact prices of dishes ordered as well as individual-specific data such as the total billing amount, date and time, table number and waiter name of each bill. From these data we use a sample of almost 90,000 customer

[18] For a step-by-step description of data collection and cleaning, we refer the reader to appendix C.1.

[19] The exact name and location is not disclosed in agreement with the restaurant's owner. Aggregation of the available data could otherwise allow local competitors to take advantage of financial figures.

choices of main dishes made at the restaurant between January 2002 and May 2009. The second data source provides first-hand information on which dishes and drinks were offered to customers at which point in time. Overall, we observe 21 different menus, offering between 24 and 30 main dishes grouped in six separate categories: traditional, fish, venison, steaks, poultry and vegetarian. In total, 75 different main dishes are included in our final choice set defined over all menus. Within each category, the available dishes are ordered according to prices, starting with the cheapest item.[20] Encoding the information from these menus, we can measure changes in the choice set that result from dish replacements across menus. As one can see from the example pages reprinted in Figure 4.2, a large number of modifications might arise: the total number of dishes within a category changes due to the addition or deletion of items, dish prices increase or fall and dishes are replaced by new items. To measure these choice set expansions, attritions and modifications, we manually construct variables that capture which choices are available and how the information is displayed to the customer. In particular, we are interested in the effect of the positioning of an alternative within a certain category on its market share. To this end, we code various dummy variables that indicate if a dish is listed at the first, last or median position within a certain category.[21]

Finally, we use official data from the State Office of Statistics as well as the CESIfo Group Munich to control for outside factors that might influence consumers' choice of products. More precisely, we use detailed regional weather data (daily information on temperature, height of precipitation and sunshine duration), as well as macroeconomic indicators (monthly unemployment rate on rural district level and an index measuring the business climate on national level).[22]

The final data set thus contains 88,113 individual choices of main dishes during the observational period from January 5^{th}, 2002 until May 29^{th}, 2009.

[20] For detailed listings of the number of observations and available dishes per menu, see Tables C.3 and C.4 in the appendix.

[21] We also constructed variables that indicate if a dish was mentioned first on a page as well as indicators for dishes that were highlighted as being a "specialty of the house" or "home-made". None of those had a significant impact on choice, such that we abstract our further analysis from them.

[22] We also included a price index to control for inflation. As results do not change significantly with inflation-adjusted prices, we will stick to the original menu prices during the analysis.

Example page from menu 7:

Fischiges

paniertes Rotbarschfilet mit Rotkohl und Kartoffeln	9,80€
Forelle „Müllerin" mit zerlassener Butter, Rotkohl und Kartoffeln	10,40€
gedünstetes Lachsfilet an grünen Nudeln mit Knoblauch-Schrimps-Soße	12,90€
naturgebratenes Pangasiusfilet (Wels) auf Zucchini-Tomaten-Bett im Reisrand	18,50€

Wildgerichte

Wildgulasch mit Rotkohl und Knödeln	10,70€
Wildschweinbraten mit Rotkohl und Klößen	11,90€
Hirschbraten mit Waldpreiselbeeren', Kroketten und Waldpilzen	12,80€
Spezialität des Hauses: Jagdherrenplatte (Wildvariationen vom Hirsch, Wildhasen und Wildschwein mit Waldpilzen und Kroketten, garniert mit Orangenscheiben)	18,90€

Example page from menu 14:

Fischiges aus Fluss und Meer

gebackene Tintenfischringe auf Salat mit Knoblauchdip und Toast	7,50€
paniertes Rotbarschfilet mit Apfelrotkohl und Kartoffeln	9,90€
Forelle „Müllerin" mit zerlassener Butter, Apfelrotkohl und Kartoffeln	10,60€
Pasta mit Riesengarnelen in einer tomatigen Knoblauch-Sahne-Soße	12,50€
gebratenes Pangasiusfilet (Wels) mit Reis, Broccoli und Dill-Sahne-Soße	12,90€

Wildspezialitäten

Wildschweinbraten mit Apfelrotkohl und Klößen	12,50€
Hirschbraten mit Waldpreiselbeeren', Kroketten und Waldpilzen	12,90€

Spezialitäten des Hauses:

Jagdherrenplatte (Wildvariationen vom Hirsch, Wildhasen und Wildschwein mit Waldpilzen und Kroketten, garniert mit Orangenscheiben)	14,80€
Bisonbraten mit Waldpilzen, Apfelrotkohl und Kroketten	17,50€

Figure 4.2: Examples of menu pages

Table 4.1 summarizes the main variables of interest as well as the most important controls.[23] The average price of a chosen main dish is EUR 10.05, with a standard deviation of about EUR 2.60. About 38% of the chosen items are located at a middle position within their menu category while only less than 25% are either on the first or last position.[24] The average individual choosing a main dish comes in a group of three or four (no_maindish= 3.62). The facts that more than half of the customers find their way to the restaurant on a weekend and almost 30% ordered while sitting at an outside table indicate that the restaurant is a day-trip destination for tourists.

Table 4.1: Summary statistics of the main explanatory variables

Variable	Mean	Std. Dev.	Min.	Max.
price_item	10.05	2.63	3.5	17.8
sum_category	4.92	1.44	2	8
first_pos	0.23	0.42	0	1
last_pos	0.22	0.42	0	1
mid_pos	0.38	0.49	0	1
adav_median	1.16	0.85	0	3.5
no_maindish	3.62	2.24	1	15
outside	0.29	0.46	0	1
business climate	96.69	7.04	82.3	108.9
evening	0.28	0.45	0	1
public holiday	0.05	0.22	0	1
weekend	0.57	0.49	0	1
N		88113		

Concerning the distribution of prices and position characteristics within the distinct menu categories (see Table 4.2), we find that the average price of venison is more than twice the price of an average vegetarian dish. Hence, price seems to be an important covariate when investigating the positional effects. Overall, variation across categories is substantial for all position dummies. Nevertheless, we do not find that individuals pick first position items, i.e. the cheapest options, more often in high price categories like venison and fish than they do in the low price segment vegetarian.[25] Finally, between 28% to 54% of the customers chose the middle posi-

[23]For an explanation of the variable names, see Table C.5 in the appendix. Moreover, a summary statistic for all variables included in the data set is given in Table C.1. We have run regressions with all covariates displayed in Table 4.1. The final specifications shown in Appendix C.3 include only the most important control variables.

[24]The middle position variable (mid_pos) is defined such that it equals the median whenever there is an unequal number of items. When there is an equal number of items within a category, the two middle items are both defined to be at the mid_pos.

[25]Note that the frequency of choosing the first or last position also depends on the overall number of items within a category. Furthermore, the cheapest vegetarian option is probably

tion from a category. This is notable since there are on average around five dishes per category. In consequence, the median and therefore compromising options have a disproportiately high market share in all categories.

Table 4.2: Summary statistics by category

		price_item		**first_pos**		**last_pos**		**mid_pos**		**adav_median**	
	N	Mean	StdDev	Mean	StdDev	Mean	StdDev	Mean	StdDev	Mean	StdDev
traditional	24,689	8.85	2.01	0.14	0.35	0.31	0.46	0.25	0.43	1.69	1.08
fish	13,568	11.07	1.64	0.19	0.40	0.18	0.38	0.37	0.48	0.99	0.62
venison	13,763	12.95	1.82	0.25	0.43	0.16	0.37	0.54	0.50	0.94	0.52
steak	19,402	10.75	1.58	0.23	0.42	0.19	0.39	0.44	0.50	0.94	0.68
poultry	7,436	9.51	2.34	0.32	0.47	0.26	0.44	0.50	0.50	0.71	0.56
vegetarian	9,255	6.36	2.00	0.42	0.49	0.17	0.38	0.28	0.45	1.16	0.67

Finally, note that the choice situations that we investigate originate from table-wise restaurant bills, i.e. in most of the cases they reflect individual choices made in the presence of a certain group.[26] Thus, group effects might influence individual choices. Due to the nature of our data set, we do not know anything about these groups except for their approximate size as measured by the total number of main dishes listed on a bill. Hence, we cannot distinguish a family from a group of colleagues or friends. Furthermore, we cannot identify who paid the bill and we do not know if the total sum was split up between the participants or if one person paid for all.[27] Therefore, we can only speculate about the effect of the presence of other people on the individual choices. There exist several papers in consumer behavior and social psychology showing that individual decisions depend upon the decisions and judgements made within a group. While most evidence indicates that group-influence leads to convergence in choices (see e.g. Levine and Moreland (1998)), a study by Ariely and Levav (2000) finds that sequential choice in group settings leads to more varied choice. The authors perform a field experiment in a Chinese restaurant and find that group presence induces individuals to make choices that produce higher variety at the group level than expected under randomization. Claiming that people's choices result from balancing different individual-alone (satisfying one's taste) versus individual-group goals (self-representation, minimizing regret

chosen very frequently because it is not only the cheapest option in that category, but also the cheapest overall.

[26] Note that only 7.13% of all observations are from a bill that contained a single main dish, our indicator for group size.

[27] Splitting up the bill is a common phenomenon in Germany, especially amongst colleagues and friends.

and information-gathering), they conclude that the observed group variety must come at the expense of personal taste satisfaction. On the other hand, Simonson and Nowlis (2000) have shown that being evaluated enhances group members' susceptibility to the compromise effect. In consequence, authors like Kivetz, Netzer, and Srinivasan (2004b) claim that group decision making will reinforce context effects. Hence, larger groups should increase the size of compromise effects instead of balancing out choice behavior. In order to control for group effects in our data, we investigate if choice patterns differ significantly for different group sizes.

4.4 Empirical Strategy

In this section, we will first state our testable hypotheses. Then, we briefly define three different descriptive measures of the compromise effect. Finally, we outline the discrete choice models used in our regressions.

4.4.1 Testable Hypotheses

As shown in Section 4.2, the compromise effect manifests itself if a choice set modification leads to a violation of the regularity condition. Moreover, even a violation of the weaker condition described as betweenness inequality confirms its existence. Hence, we test for the following correlation between choice set modification and choice probabilities:

Hypothesis 1 *If a dish moves from an extreme to an intermediate position within a menu category, then its* absolute *market share increases (violation of the regularity condition).*

Hypothesis 2 *If a dish moves from an extreme to an intermediate position within a menu category, then its market share* relative *to the other option(s) in the original choice set increases (violation of the betweenness inequality).*

To test whether these conditions are violated in our data, we proceed in two steps. First, we present changes in market shares for all pure choice set expansions available in our data. Second, we proceed by modeling individual choice behavior using discrete choice models that allow us to test for position effects using all of the available variation in our data.

4.4.2 Descriptive Measures

In order to test the above hypotheses, we will identify all cases of pure choice set expansions, i.e. all situations were the introduction of a new menu increases the number of dishes within a category by exactly one. Moreover, the added alternative has to be on the upper or lower boundary of the price range, such that one of the existing options moves from being an extreme to being a compromise. Besides the addition of such an extreme choice option, no other characteristic is allowed to change. In other words, prices and positions of all other alternatives have to stay constant. Overall, we find nine pure cases in our data set. In three out of these cases, a low-price option is added. Investigating each case separately, we will compare market shares before and after the menu change. Total demand in each case is set equal to the number of individuals having chosen a dish from the respective category, such that the shares of all alternatives in the respective category add up to 1.

Our first measure is the difference in unconditional *absolute* market shares for the option that switches from being the extreme to being a more moderate choice: $P_x(i) - P(i)$ (=compromise effect measure I). Here, i refers to the option that we are investigating, whereas the subscript x indicates that the additional new option is present in the choice set. According to Hypothesis 1, this difference should be larger than zero, indicating a violation of regularity.

The second indicator measures the change in the market share of the option of interest *relative* to the option at the opposite end of the ascending-price ordering. It is defined as $P_x(i, \{i, j\}) - P(i, \{i, j\})$ (=compromise effect measure II), with j being equal to the lowest-price option in the category if a new highest-price option is added and j being equal to the highest-price option if the new alternative enters the list at the low-price end. In line with Hypothesis 2, this difference should be larger than zero, indicating a violation of the betweenness inequality.

Finally, we introduce a third relative measure to account for the fact that most of our choice sets contain more than three alternatives.[28] Instead of comparing the market share variation relative to one other option in the original set, we now measure the difference in the market share of the option of interest relative to the complete original choice set before and after the set expansion. Hence, we compute $P_x(i|\{.\}) - P(i|\{.\})$ (=compromise effect measure III) with x being the alternative that is added to the choice set after the menu change and $\{.\}$ containing all options available in both menus. Once again, we expect the difference to be positive in line with Hypothesis 2.

4.4.3 Regression Analysis

In order to control for other alternative-specific as well as individual-specific factors that might have influenced the choice of our individual decision makers, we run a conditional logit model for each of the nine pure cases.[29] Let us assume that, for individual i and dish choice k, the utility U_{ik} is additively separable in a deterministic component V_{ik} and a random component ϵ_{ik}:

$$U_{ik} = V_{ik} + \epsilon_{ik} \qquad \forall \quad k = \{1, ..., K\}$$

In our data, we only observe the choice that yields the highest utility:

$$\begin{aligned} Pr(y_i = k) &= Pr\left(U_{ik} \geq U_{il}\right), \quad \forall \quad l \neq k \in K, \\ &= Pr(\epsilon_{il} - \epsilon_{ik} \leq V_{ik} - V_{il}), \quad \forall \quad l \neq k \in K, \end{aligned}$$

[28] This is in sharp contrast to most of the lab experiments. In these artificial setups, the original choice set consists only of two alternatives. Hence, changes in the relative market share can only be measured with respect to the one other alternative that stays at the extreme end of the price-quality trade-off line.

[29] See, for example, Train (2003) or Cameron and Trivedi (2009) for detailed descriptions of the following discrete choice models. Note that the estimation of a conditional logit might seem controversial at first sight as this model explicitly assumes that the IIA is fulfilled. However, this assumption has to hold true only conditional on the included regressors. Hence, controlling for compromise effects measures is vital for the IIA assumption to hold.

where the deterministic part of this utility is explained by alternative specific regressors (X_{ik}) such as the price and case-specific regressors (Z_i) such as the table size, the weather or the overall economic situation:

$$V_{ik} = X'_{ik}\beta + Z'_i\gamma_k$$

X also contains an alternative-specific intercept that captures time-invariant, unexplained utility similar to a fixed effect. Finally, according to the compromise effect theory, an individual's utility from a certain choice will be influenced by the addition of a new "extreme-position" dish. The introduction of this item will be measured by a dummy indicating a change of menus. Hence, the utility specification for each of the nine pure cases is given as follows:

$$U_{ik} = \quad \beta_0 + \beta_1 \ price\_item_{ik} + \gamma_{1k} \ menu\_change_i + Z'_i\gamma_k + \epsilon_{ik} \tag{4.5}$$

The probability of choosing k conditional on all observables can then be written as:

$$p_{ik} = Pr(y_i = k|X_{ik}, \beta, Z_i, \gamma_k) = F_k(X_{ik}, \beta, Z_i, \gamma_k)$$

and the individual likelihood contribution follows to be:

$$L_i(y_i|X_{ik}, \beta, Z_i, \gamma_k) = \prod_{k=1}^{K} p_{ik}$$

Under the assumption that, conditional on the position in the menu and other observables, independence of irrelevant alternatives holds, we can write choice probabilities as:

$$p_{ik} = \frac{exp(X'_{ik}\beta + Z'_i\gamma_k)}{\sum_{l=1}^{m} exp(X'_{il}\beta + Z'_i\gamma_l)}$$

Turning to the coefficients' interpretation, note first that the alternative-specific intercepts reflect the desirability of each alternative due to its unmeasured time-invariant attributes. The coefficients of alternative-specific regressors, such as the price of a dish, can be interpreted very easily. A positive coefficient means that if the respective regressor increases for one of the choice options, then that item is chosen more often and all other dishes together are chosen less often. Case-specific regressors, on the other hand, are to be interpreted as parameters of a binary logit model against the base alternative. In our pure cases, this base alternative will correspond to option j as described in 4.4.2. Hence, relative to the probability of the base alternative, an increase of the regressor leads to an increase in the choice probability of the alternative under consideration if the respective coefficient has a positive sign. In order to give a meaning to the choice probability of the compromise option, we calculate the marginal effect at the mean of the "menu_change" dummy. We thus test the effect of a one-unit increase in the respective regressor on the choice probability of an option for a fictional decision maker, while keeping all other variables at their mean values.

As these nine pure cases correspond only to a small fraction (29.3%) of the complete data set, we run additional regressions in order to exploit all menu changes, even if they induce multiple and simultaneous modifications of the category-wise choice sets. To this end, we coded several alternative-specific dummies (see Section 4.3 for details), indicating the location of an alternative within the respective category. In particular, we are interested in the coefficients of the following exogenous variables: first_pos, last_pos, mid_pos and adav_median. Under the above hypotheses, all signs except for the one of mid_pos should be negative.[30]

We proceed with these full sample regressions in two steps: First, we estimate conditional logits as well as mixed logits for all menu categories separately. This allows us to find out whether or not the compromise effect is present in all distinct menu categories. Second, we use a random draw of the complete menu including all six categories and run both conditional logit as well as mixed logit regressions.[31][32] Using the entire sample, we can also control for the possibility that changing the choice set within a certain category might lead customers to switch to a completely

[30]The utility specification described in equation (4.5) will thus change to be equal to

$$U_{ik} = \quad \beta_0 + \beta_1 \; price\_item_{ik} + \beta_2 \; position \; dummy_{ik} + Z_i' \gamma_k + \epsilon_{ik}$$

[31]Due to computational restrictions, we reduce the sample size by randomly drawing a subsample of observations. As we have 75 alternatives in the full menu sample, the simulations would otherwise take too long. In both steps, we have to estimate the models using an unbalanced choice set as the number of available alternatives varies across individual observations.

[32]In order to reduce computation time, we used the bwGRiD (http://www.bw-grid.de/) cluster server MA/HD for our full menu sample estimations.

different category.[33]

Estimating mixed logits in both steps enables us to further relax the IIA assumption by introducing normally distributed coefficients on price and item position, i.e. $\beta_i = \beta + v_i$ (with $v_i \sim N(\mathbf{0}, \Sigma_\beta)$), thereby introducing an error correlation across choices. Conditional on the unobservable random part (v_i), choice probabilities are thus given by:

$$p_{ik}|v_i = \frac{exp(X'_{ik}\beta + X'_{ik}v_i + Z'_i\gamma_k)}{\sum_{l=1}^{m} exp(X'_{il}\beta + X'_{il}v_i + Z'_i\gamma_l)} \tag{4.6}$$

The unobserved component v_i is integrated out by numerical simulation using a

sequence of Halton draws to simulate the probabilities.[34]

We conclude this section with a final remark concerning the possible endogeneity of the price variable. Oftentimes, it is necessary for prices to be instrumented in structural demand analysis. In our case, we think this is not necessary. First, we use individual choice data rather than aggregated demand observations. Second, we know from data inspection (see Section 4.3 and Appendix C.1) and interviews with the owner that the restaurant mostly serves occasional walk-in customers, while the number of regular customers, who might think that their individual demand is able to influence the composition of menus, is very small. Therefore, we assume that consumers take menu prices as given when choosing their dish. As far as the restaurant's pricing strategy is concerned, we know from several interviews with the owner that prices of the restaurant are calculated exclusively based on a cost-plus-markup method. Besides, over the year, the menu is changed mainly in response to a change of the season. Seasonal changes of menu items, however, are clearly exogenous to individual customer choice. [35]

[33]Unfortunately, we cannot separately test for the effect of being the cheapest (or most expensive) item within categories and overall, because the cheapest and most expensive items remain in the same categories over time.

[34]Halton draws are a way of systematic sampling that ensures a good coverage of the underlying distribution even with a limited number of draws.

[35]Unfortunately, we cannot test this assumption because we have no data on input costs for each dish or on local competitors' prices for similar dishes that might serve as valid instruments for prices.

4.5 Results

In the following, we will first present the descriptive measures as well as the regression results for the nine pure cases of choice set expansion before turning to the full sample analysis.

4.5.1 Pure Cases

Tables 4.3 and 4.4 present the descriptive measures as defined in the previous section. Table 4.3 shows that all cases except for case 4 provide evidence for the existence of a compromise effect. First, the difference in absolute market shares $P_x(i) - P(i)$ for the compromise option (printed in bold letters) is positive for eight out of the nine cases. Due to the rather small increases (between .92 and 7.78 percentage points), however, only one of the differences is significant. Turning to the change in market share relative to the most extreme alternative in the original set, $P_x(i, \{i, j\}) - P(i, \{i, j\})$, we find that all choice set expansions except for case 4 lead to a violation of the betweenness inequality. In line with the fact that this measure refers to the rationality condition that is more easily violated, five of the measures are significantly positive. This result is confirmed by our third compromise effect measure $P_x(i|\{.\}) - P(i|\{.\})$. Overall, we conclude that there is evidence for the compromise effect, although the changes in market shares are quite small. Two other things are worth noticing: First, the effect is more significant in small choice sets (cases 5, 6 and 7), a result which is consistent with previous findings in lab experiments (see e.g. Kivetz, Netzer, and Srinivasan (2004a)).[36] Second, turning to Table 4.4, we see that the compromise effect is not necessarily the consequence of a substitution away from the extreme option to the now more compromising alternative. Instead, relative market shares indicate that customers might also switch away from other options in the choice set. Therefore, it seems useful to investigate the robustness of the effect by controlling for alternative-

[36]Note that these three cases all correspond to a set expansion by addition of a new highest-price item. Thus, we cannot say if the addition of a new cheapest-price item would also lead to more significant measures in smaller choice sets.

specific constants as well as other case-specific control variables.[37]

Table 4.3: absolute market shares and descriptive compromise effect measures I and II

case 1			case 2			case 3		
position	menu I	menu II	position	menu I	menu II	position	menu I	menu II
A		12.03	A	8.68	6.47	A		11.09
B	**7.65**	**8.58**	B	37.19	33.81	**B**	**23.71**	**25.50**
C	22.06	15.08	C	28.51	20.14	C	38.26	26.16
D	28.83	25.73	D	7.02	6.47	D	19.69	23.01
E	15.66	15.35	**E**	**18.60**	**20.14**	E	18.34	14.24
F	25.80	23.24	F		12.95			
N obs.	562	723	N obs.	242	139	N obs.	447	604
Pearson chi2(5) = 1.2e+03 Pr = 0.000			Pearson chi2(5) = 173.7027 Pr = 0.000			Pearson chi2(4) = 437.1481 Pr = 0.000		
$P_A(B) - P(B) = 0.92$			$P_F(E) - P(E) = 1.55$			$P_A(B) - P(B) = 1.78$		
$P_A(B, \{B, F\}) - P(B, \{B, F\}) = 4.08$			$P_F(E, \{A, E\}) - P(E, \{A, E\}) = 7.49$			$P_A(B, \{B, E\}) - P(B, \{B, E\}) = 7.78^{*}$		

case 4			case 5			case 6		
position	menu I	menu II	position	menu I	menu II	position	menu I	menu II
A		13.38	A	26.66	24.84	A	25.00	8.54
B	*20.82*	*16.56*	B	46.26	36.87	B	55.73	25.20
C	32.39	30.57	**C**	**27.08**	**28.77**	**C**	**19.27**	**24.39**
D	30.85	28.03	D		9.52	D		41.87
E	15.94	11.46						
N obs.	389	157	N obs.	4,265	5,367	N obs.	192	492
Pearson chi2(4) = 265.6338 Pr = 0.000			Pearson chi2(3) =1.4e+03 Pr = 0.000			Pearson chi2(3) = 445.5654 Pr = 0.000		
$P_A(B) - P(B) = -4.26$			$P_D(C) - P(C) = 1.69^{**}$			$P_D(C) - P(C) = 5.12^{*}$		
$P_A(B, \{B, E\}) - P(B, \{B, E\}) = -2.45$			$P_D(C, \{A, C\}) - P(C, \{A, C\}) = 3.27^{***}$			$P_D(C, \{A, C\}) - P(C, \{A, C\}) = 30.54^{***}$		

case 7			case 8			case 9		
position	menu I	menu II	position	menu I	menu II	position	menu I	menu II
A	27.68	17.20	A	45.00	48.06	A	44.79	36.71
B	50.87	38.53	B	16.50	17.95	B	17.35	17.72
C	**21.45**	**23.39**	C	28.50	10.17	C	13.88	12.03
D		20.87	**D**	**10.00**	**12.76**	D	12.62	5.70
			E		11.07	**E**	**11.36**	**14.56**
						F		13.29
N obs.	289	436	N obs.	200	1,003	N obs.	317	158
Pearson chi2(3) = 226.8622 Pr = 0.000			Pearson chi2(4) = 264.3349 Pr = 0.000			Pearson chi2(5) = 249.0974 Pr = 0.000		
$P_D(C) - P(C) = 1.94$			$P_E(D) - P(D) = 2.76$			$P_F(E) - P(E) = 3.20$		
$P_D(C, \{A, C\}) - P(C, \{A, C\}) = 13.97^{***}$			$P_E(D, \{A, D\}) - P(D, \{A, D\}) = 2.80$			$P_F(E, \{A, E\}) - P(E, \{A, E\}) = 8.17^{*}$		

All market shares are given as percentages. Differences between absolute or relative market shares are in percentage points.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Significance levels refer to one-sided t-tests (H_0: difference in shares < 0).
Source: Own Data, 2002-2009.

Table C.6 in the appendix shows the results of a conditional logit regression with intercepts, but excluding further case-specific controls. The coefficients displayed for the case-specific regressor "menu_change" refer to the coefficient of the com-

[37]The case-specific control variables that are included in all of the following regressions were chosen after thorough inspection of all possible controls. In order to reduce model complexity, we only included business climate, weekend, public holiday, evening and no_maindish as these were the ones with significant coefficients for a wide range of alternatives. All other tested control variables (see Table C.1 for a complete list of all case-specific variables) proved insignificant for the most part and their inclusion in the model does not affect the results.

Table 4.4: relative market shares and descriptive compromise effect measure III

	case 1				case 2		
position	menu I $P(i\|\{B,C,D,E,F\})$	menu II $P_A(i\|\{B,C,D,E,F\})$	Diff	position	menu I $P(i\|\{A,B,C,D,E\})$	menu II $P_F(i\|\{A,B,C,D,E\})$	Diff
A				A	0.0868	0.0744	-.0124 (0.6565)
B	**0.0765**	**0.0975**	**.0210 (0.1003)**	B	0.3719	0.3884	.0165 (0.3801)
C	0.2206	0.1714	-.0493 (0.9842)	C	0.2851	0.2314	-.0537 (0.8616)
D	0.2883	0.2925	.0042 (0.4366)	D	0.0702	0.0744	.0041 (0.4430)
E	0.1566	0.1745	.0179 (0.2027)	**E**	**0.1860**	**0.2314**	**.0455 (0.1549)**
F	0.2580	0.2642	.0061 (0.4047)	F			
$\sum$	1.0000	1.0000		$\sum$	1.0000	1.0000	

	case 3				case 4		
position	menu I $P(i\|\{B,C,D,E\})$	menu II $P_A(i\|\{B,C,D,E\})$	Diff	position	menu I $P(i\|\{B,C,D,E\})$	menu II $P_A(i\|\{B,C,D,E\})$	Diff
A				A			
B	**0.2371**	**0.2868**	**.0496** (0.0394)**	**B**	*0.2082*	*0.1912*	*-.0170 (0.6642)*
C	0.3826	0.2942	-.0883 (0.9983)	C	0.3239	0.3529	.0290 (0.2684)
D	0.1969	0.2588	.0620** (0.0108)	D	0.3085	0.3235	.0150 (0.3725)
E	0.1834	0.1601	-.0233 (0.8330)	E	0.1594	0.1324	-.0270 (0.7744)
$\sum$	1.0000	1.0000		$\sum$	1.0000	1.0000	

	case 5				case 6		
position	menu I $P(i\|\{A,B,C\})$	menu II $P_D(i\|\{A,B,C\})$	Diff	position	menu I $P(i\|\{A,B,C\})$	menu II $P_D(i\|\{A,B,C\})$	Diff
A	0.2666	0.2745	.0079 (0.1980)	A	0.2500	0.1469	-.1031 (0.9977)
B	0.4626	0.4075	-.0551*** (0.0000)	B	0.5573	0.4336	-.1237 (0.9961)
C	**0.2708**	**0.3180**	**.0471*** (0.0000)**	**C**	**0.1927**	**0.4196**	**.2269*** (0.0000)**
D				D			
$\sum$	1.0000	1.0000		$\sum$	1.0000	1.0000	

	case 7				case 8		
position	menu I $P(i\|\{A,B,C\})$	menu II $P_D(i\|\{A,B,C\})$	Diff	position	menu I $P(i\|\{A,B,C,D\})$	menu II $P_E(i\|\{A,B,C,D\})$	Diff
A	0.2768	0.2174	-.0594 (0.9584)	A	0.4500	0.5404	.0904** (0.0104)
B	0.5087	0.4870	-.0217 (0.7065)	B	0.1650	0.2018	.0368 (0.1178)
C	**0.2145**	**0.2957**	**.0811** (0.0101)**	C	0.2850	0.1143	-.1707 (1.0000)
D				**D**	**0.1000**	**0.1435**	**.0435* (0.0523)**
$\sum$	1.0000	1.0000		E			
				$\sum$	1.0000	1.0000	

	case 9		
position	menu I $P(i\|\{A,B,C,D,E\})$	menu II $P_F(i\|\{A,B,C,D,E\})$	Diff
A	0.4479	0.4234	-.0246 (0.6855)
B	0.1735	0.2044	.0309 (0.2179)
C	0.1388	0.1387	-.0001 (0.5013)
D	0.1262	0.0657	-.0605 (0.9716)
E	**0.1136**	**0.1679**	**.0543* (0.0573)**
F			
$\sum$	1.0000	1.0000	

Diff= $P_x(i|\{.\}) - P(i|\{.\})$ with x being the alternative that is added to the choice set in menu II.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Significance levels refer to one-sided t-tests (H_0: Diff > 0). p-values in parentheses.

Source: Own Data, 2002-2009.

promising choice option. The same applies to the presented marginal effect at the mean. In accordance with our hypotheses, we expect a positive sign. In line with the descriptives, our hypotheses are mostly confirmed, although only the coefficients for cases with small choice sets are significant. Taking case 6, for example, we see that the menu change/choice set expansion leads to a highly significant increase in alternative C's market share in comparison to the market share of alternative A by 1.31. Including case-specific controls that account for the individual purchase background changes the size of the coefficients, but not their significance (see Table C.7 in the appendix). Inspecting the likelihoods with and without additional controls as well as the chi-square statistic of the likelihood ratio test for the joint significance of coefficients ($Prob > chi2$), we conclude that the second specification is preferred.

4.5.2 Complete Sample Evidence

Turning to the regression results for the full data set, we first present category-wise estimations and then turn to the estimations over all alternatives. Tables C.8 - C.13 show different conditional logit specifications, while Tables C.14 - C.19 present mixed logit estimations for the same variable specifications. First, we find that the price has a significantly negative coefficient for all categories except vegetarian both in all conditional logit and all mixed logit specifications. A reason for the positive price coefficient in the vegetarian category might be that it is the cheapest category within the menu. Hence, once people have decided to go for a vegetarian dish, they might be less price sensitive knowing that they already chose a rather cheap option as compared to the other categories' price levels. Turning to the dummy variables capturing the existence of a compromise effect, we conclude that in both multinomial models, first_pos and last_pos have a significantly negative sign in most categories. This means that being an extreme option within a certain category decreases the choice probability of the respective alternative significantly. Being the low-price, low-quality item within a category has a positive impact on the probability of being chosen only in category 3 (venison). A possible explanation for this exception might once again be the overall price level of the category. As can be seen in Table 4.2, venison is the menu category with the highest average price level.[38] In consequence, customers might pick the lowest price alternative, because they know that it is still a medium-price, medium-quality option with regard to the complete menu. Another unexpected sign is found for the dummy last_pos in category 1 (traditional dishes). Unfortunately, we do not have an explanation for this result.[39] The coefficients of the dummy mid_pos that indicates if a dish is located at the median position within a category are in general positive and significant, confirming hypothesis 1. The conditional logit estimates for category 3 as well as the mixed logit estimates for categories 1 and 6, however, are insignificant. Finally, the variable "adav_median" measures the absolute distance from the median position. The coefficients in the conditional logits are significantly negative (as implied by the compromise effect) only for category 2, 4 and 5, while they are positive but insignificant for the remaining menu categories. Looking at the mixed logit regressions, the significantly negative coefficients for the former categories are confirmed. For the three remaining categories, the respective coefficients are insignificant both with and without controls. Summing up, our category-wise regressions support the evidence found in the nine pure cases. A compromise effect seems to be present in all of the categories. Yet, we find that overall price levels also matter.

We conclude the analysis by estimating a set of full menu multinomial models. As already said in the data section, there are 75 different alternatives that were offered over the course of the observational period. Out of these alternatives, about a third was listed in each of the respective menus. Hence, we use an unbalanced choice set in the estimations. In order to keep running times for the computations tractable, we included only alternative-specific constants and no_maindish as controls. Tables C.20 and C.21 show the results. The price coefficient is significantly negative in all specifications. Turning to the position dummies, however, we find that the only significant effect is found for last_pos. Hence, if a dish is the highest-price, highest-

[38] We also tested for the influence of category-wise price levels on consumer choices using a two-level nested logit model with prices as explanatory variable on the first-level. Category price coefficients were not significant. This shows that the beliefs of consumers about the average price of a respective category are fixed and not reestimated at every restaurant visit.

[39] Our hypothesis that the large variety in dishes subsumed under category 1 would explain this anomaly was not confirmed in robustness checks. More precisely, category 1 includes a variety of dishes ranging from simple and low-price soups to homemade roasts. We included respective dummies (homemade, soup) in our regressions to control for these exceptions. In contrast to our hypothesis, however, these dummies were neither significant nor did they have any impact on the coefficient of last_pos.

quality option of a category, this reduces its choice probability. In contrast, being the lowest-price, lowest-quality option as well as being at the compromising middle position does not have a significant impact on market shares. This stands in contrast to experimental findings indicating that the low-quality, low-price alternative is relatively less popular in choice sets with a price-quality trade-off, while this is not true for the high quality, high price alternative (see e.g. Simonson and Tversky (1992)). In consequence, our customers might be seen as more extremeness averse in the price dimension than they are with respect to dish quality.[40] In a recent experimental study with highly different product segments, Gierl and Stiegelmayr (2010) confirm the result by Simonson and Tversky (1992) for high price product categories. Yet, they also show that the opposite holds true for products with low absolute price levels. In these latter segments, the high-quality, high-price option within the choice set looses the most after introducing a compromise option - a result which confirms our findings. Turning to the two remaining regressors, our regressions for the entire sample provide mixed evidence. While mid_pos is significantly positive in the conditional logit model, it is insignificant in the mixed logit model. On the other hand, adav_median has a significantly negative impact on choice probabilities according to the mixed logit model, while its coefficient in the conditional logit is insignificant. Turning to model fit, we see that the mixed logit specifications have substantially higher log likelihoods than the conditional logits. Moreover, the fact that at least some of the alternative-specific regressors of the mixed logits have a standard deviation that is significantly different from zero indicates that we should prefer the more flexible random parameters model to the conditional logit.

4.6 Conclusion

In this paper, we provide evidence for the existence of a compromise effect in natural choice situations by means of a newly generated data set comprising more than 88,000 individual choices of main courses from 21 different restaurant menus

[40]This implication holds true if we believe Tversky and Simonson (1993)'s explanation for the compromise effect.

offered over a period of more than 7 years. Two sets of results are presented: First, we inspect all menu changes where an extreme option within a main dish category becomes a compromising option and investigate the existence of compromise effects by means of the regularity condition and betweenness inequality. Second, we conduct regression analyses for each sub-category of menu items, as well as for the entire list of main dishes on the menu. To this end, we estimate conditional and mixed logit models, where in the latter we allow for heterogeneous coefficients and correlated unobserved heterogeneity across choices.

We find that the compromise effect prevails in both types of analysis as well as in all sub-categories of main dishes. The effect, however, is stronger if an item moves from being an item at the top end, rather than the bottom end, of the price spectrum to being a compromise option. Furthermore, our results are robust to controlling for weather effects, business cycle climate as well as the weekday or time. Yet, controlling for confounding factors such as the background context of individual decision makers is important, as it can change the size of the compromise effect substantially. Besides, group size seems to matter in some specifications although the direction of the effect is ambivalent. This paper is thus the first to provide "real-world" evidence for the existence of the compromise effect that may be valuable both to economic theorists and practitioners.

The results of this paper are relevant for future research purposes, applications and policy making. First, theoretical researchers might want to rely on them when building models for economic choice behavior, whenever compromise effects are of importance. Second, our results are relevant to natural settings where choice sets are of strategic importance, such as in marketing or Internet applications. In fact, choices from ordered lists, where consumers have to weight alternatives with respect to several choice attributes are omnipresent. Examples are Internet shops where alternatives are listed with respect to price or product characteristics, Likert scales in surveys or even ballot papers where parties are ordered with respect to their (left-to-right) political orientation. Third, for policy makers and regulators, it is important to know to what extent firms can deceive customers by introducing very extreme, overpriced options to make other menu items seem more attractive.

Nevertheless, the present study faces several limitations. First, our data set does not contain customer-specific information. Hence, although unlikely a problem in our setting, we are not able to control for a changing customer base across menu periods. Second, any variation on extreme and compromise options that we exploit comes exclusively from the 21 menu changes observed in the data. Thus, the variability in the data is rather low, which leads to numerical problems in the estimation of sophisticated models with many parameters. Third, although the customers in the data are price-takers, we cannot completely rule out that menus changed in response to previous customer decisions. This could potentially introduce spurious correlation, although it is not clear why this should influence compromise options differently from extreme options. Finally, all our evidence is generated using data on one specific restaurant and results may differ in different choice settings or locations. Sheng, Parker, and Nakamoto (2005), for example, investigate the effect of certain consumer characteristics on the compromise effect. They focus on product familiarity and attribute-importance structure. Furthermore, they show that other factors such as the availability and quantity of information on the product as well as the product category itself might have an influence. Hence, the compromise effect may not necessarily prove valid in different settings or for different agents.

Future research is needed in two respects. First, the present analysis should be replicated in different settings, countries and sales environments to see if our results persist among a different customer base. Second, there is a need for data analysis on natural choice situations for which customer-specific information is available. Such data could provide answers as to which types of customers are particularly sensitive to the compromise effect and provide valuable information for practitioners, regulators and policy makers.

Appendix A

Appendix of Chapter 2

A.1 Partial Derivatives of π_{ASE}^D

$$\frac{\partial \pi_{ASE}^D}{\partial \Delta} = \begin{cases} \frac{\Delta}{8(t+a)} & \text{if } k < \frac{7}{2}(t-a) - \frac{\Delta(t-a)}{2(t+a)} \\ \frac{-(2k-7(t-a)-\Delta)}{16t} & \text{if } \frac{7}{2}(t-a) - \frac{\Delta(t-a)}{2(t+a)} \le k < \frac{7}{2}(t-a) + \frac{\Delta}{2} \\ 0 & \text{if } \frac{7}{2}(t-a) + \frac{\Delta}{2} \le k \end{cases}$$

$$\frac{\partial \pi_{ASE}^D}{\partial a} = \begin{cases} \frac{1}{16}\left(\frac{4k^2-(t-a)^2}{(t-a)^2} - \frac{\Delta^2}{(t+a)^2}\right) & \text{if } k < \frac{7}{2}(t-a) - \frac{\Delta(t-a)}{2(t+a)} \\ \frac{14k-t+49a-7\Delta}{16t} & \text{if } \frac{7}{2}(t-a) - \frac{\Delta(t-a)}{2(t+a)} \le k < \frac{7}{2}(t-a) + \frac{\Delta}{2} \\ 3 & \text{if } \frac{7}{2}(t-a) + \frac{\Delta}{2} \le k \end{cases}$$

$$\frac{\partial \pi_{ASE}^D}{\partial t} = \begin{cases} \frac{1}{16}\left(\frac{-(4k^2-(t-a)^2)}{(t-a)^2} - \frac{\Delta^2}{(t+a)^2}\right) & \text{if } k < \frac{7}{2}(t-a) - \frac{\Delta(t-a)}{2(t+a)} \\ -\frac{(2k+7a-7\Delta)^2}{32t^2} - \frac{47}{32} & \text{if } \frac{7}{2}(t-a) - \frac{\Delta(t-a)}{2(t+a)} \le k < \frac{7}{2}(t-a) + \frac{\Delta}{2} \\ -3 & \text{if } \frac{7}{2}(t-a) + \frac{\Delta}{2} \le k \end{cases}$$

$$\frac{\partial \pi_{ASE}^D}{\partial k} = \begin{cases} \frac{k}{2(t-a)} + \frac{1}{4} & \text{if } k < \frac{7}{2}(t-a) - \frac{\Delta(t-a)}{2(t+a)} \\ \frac{2k+9t+7a-\Delta}{8t} & \text{if } \frac{7}{2}(t-a) - \frac{\Delta(t-a)}{2(t+a)} \le k < \frac{7}{2}(t-a) + \frac{\Delta}{2} \\ 2 & \text{if } \frac{7}{2}(t-a) + \frac{\Delta}{2} \le k \end{cases}$$

It is easy to show that $\frac{\partial \pi_{ASE}^D}{\partial \Delta}, \frac{\partial \pi_{ASE}^D}{\partial a}, \frac{\partial \pi_{ASE}^D}{\partial k} \ge 0$ and $\frac{\partial \pi_{ASE}^D}{\partial t} \le 0$ given that $a, t, \Delta > 0$ by assumption and assumptions 2.1 and 2.2 imply that $t > a$ as well as $k \ge$

$\frac{3}{2}(t-a)+\frac{\Delta}{2}$. Furthermore, k has to lie between the thresholds defining the respective part of the deviation profit.

A.2 Partial Derivatives of $\pi^D_{ASE} - \pi^D_{SE}$

$$\frac{\partial(\pi^D_{ASE} - \pi^D_{SE})}{\partial\Delta} = \begin{cases} \frac{\Delta}{8(t+a)} & \text{if } k < \frac{7}{2}(t-a) - \frac{\Delta(t-a)}{2(t+a)} \\ \frac{-(2k-7(t-a)-\Delta)}{16t} & \text{if } \frac{7}{2}(t-a) - \frac{\Delta(t-a)}{2(t+a)} \leq k < \frac{7}{2}(t-a) + \frac{\Delta}{2} \\ 0 & \text{if } \frac{7}{2}(t-a) + \frac{\Delta}{2} \leq k \end{cases}$$

$$\frac{\partial(\pi^D_{ASE} - \pi^D_{SE})}{\partial a} = \begin{cases} -\frac{\Delta^2}{16(t+a)^2} & \text{if } k < \frac{7}{2}(t-a) - \frac{\Delta(t-a)}{2(t+a)} \\ -\frac{k^2}{4(t-a)^2} + \frac{7(7a-\Delta+2k)}{16t} & \text{if } \frac{7}{2}(t-a) - \frac{\Delta(t-a)}{2(t+a)} \leq k < \frac{7}{2}(t-a) \\ \frac{7(2k-7(t-a)-\Delta)}{16t} & \text{if } \frac{7}{2}(t-a) \leq k < \frac{7}{2}(t-a) + \frac{\Delta}{2} \\ 0 & \text{if } \frac{7}{2}(t-a) + \frac{\Delta}{2} \leq k \end{cases}$$

$$\frac{\partial(\pi^D_{ASE} - \pi^D_{SE})}{\partial t} = \begin{cases} -\frac{\Delta^2}{16(t+a)^2} & \text{if } k < \frac{7}{2}(t-a) - \frac{\Delta(t-a)}{2(t+a)} \\ \frac{1}{32}(\frac{8k^2}{(t-a)^2} - \frac{(2k+7a-\Delta)^2}{t^2} - 49) & \text{if } \frac{7(t-a)}{2} - \frac{\Delta(t-a)}{2(t+a)} \leq k < \frac{7(t-a)}{2} \\ \frac{1}{32}(49 - \frac{(2k+7a-\Delta)^2}{t^2}) & \text{if } \frac{7}{2}(t-a) \leq k < \frac{7}{2}(t-a) + \frac{\Delta}{2} \\ 0 & \text{if } \frac{7}{2}(t-a) + \frac{\Delta}{2} \leq k \end{cases}$$

$$\frac{\partial(\pi^D_{ASE} - \pi^D_{SE})}{\partial k} = \begin{cases} 0 & \text{if } k < \frac{7}{2}(t-a) - \frac{\Delta(t-a)}{2(t+a)} \\ \frac{1}{8}(7 - \frac{4k}{t-a} + \frac{7a-\Delta+2k}{t}) & \text{if } \frac{7}{2}(t-a) - \frac{\Delta(t-a)}{2(t+a)} \leq k < \frac{7}{2}(t-a) \\ \frac{2k-7(t-a)-\Delta}{8t} & \text{if } \frac{7}{2}(t-a) \leq k < \frac{7}{2}(t-a) + \frac{\Delta}{2} \\ 0 & \text{if } \frac{7}{2}(t-a) + \frac{\Delta}{2} \leq k \end{cases}$$

It is easy to show that $\frac{\partial(\pi^D_{ASE}-\pi^D_{SE})}{\partial\Delta} \geq 0$, $\frac{\partial(\pi^D_{ASE}-\pi^D_{SE})}{\partial a} \leq 0$, $\frac{\partial(\pi^D_{ASE}-\pi^D_{SE})}{\partial k} \leq 0$ and $\frac{\partial(\pi^D_{ASE}-\pi^D_{SE})}{\partial t} \lesseqgtr 0$ given that $a, t, \Delta > 0$, $t > a$ as well as $k \geq \frac{3}{2}(t-a) + \frac{\Delta}{2}$ by assumption. Furthermore, k has to lie between the thresholds defining the respective part of the difference in deviation profits.

A.3 Critical Discount Factor for Collusion only on Side 2

If $\Delta < t + a$:

$$\hat{\delta}_{OC2} = \begin{cases} \text{(i)} \; \frac{(2k-3(t-a)+\Delta)(t^2-a^2+\Delta^2)}{5a^3+\Delta^3+\Delta^2(2k-3t)+9\Delta t^2+t^2(2k+5t)-a^2(9\Delta+2k+5t)+a(3\Delta^2-5t^2)} \\ \qquad \text{if } k < \frac{7}{2}t - \frac{3}{2}a - \frac{\Delta}{2} - \frac{4a\Delta t}{t^2-a^2+a\Delta} \\ \text{(ii)} \; \frac{\left[\begin{array}{c}(a-\Delta)^2(3a+\Delta+2k)^2 - 2(7a-3\Delta)(a-\Delta)(3a+\Delta+2k)t \\ +(7a-3\Delta)^2t^2 + 8(3a+\Delta+2k)t^3 - 40t^4\end{array}\right]}{\left[\begin{array}{c}(a-\Delta)^2(3a+\Delta+2k)^2 - 2(7a-3\Delta)(a-\Delta)(3a+\Delta+2k)t \\ +(25a^2 - 2a(13\Delta+8k) + \Delta(17\Delta+16k))t^2 + 8(9a-\Delta+4k)t^3 - 64t^4\end{array}\right]} \\ \qquad \text{if } \frac{7}{2}t - \frac{3}{2}a - \frac{\Delta}{2} - \frac{4a\Delta t}{t^2-a^2+a\Delta} \leq k \text{ and} \\ \quad \left[\left(k < \frac{-3a^2+2a\Delta+\Delta^2+7at-3\Delta t+4t^2}{2(a-\Delta)} \wedge \Delta < a\right) \vee \left(k < \frac{(3a+\Delta-4t)(t-a+\Delta)}{2(a-\Delta)} \wedge \Delta \geq a\right)\right] \\ \text{(iii)} \; \frac{(a-\Delta)(3a+\Delta+2k)+2(-2a+2\Delta+k)t-7t^2}{2(a+\Delta+2k-5t)t)} \quad \text{if } \frac{-3a^2+2a\Delta+\Delta^2+7at-3\Delta t+4t^2}{2(a-\Delta)} \leq k \wedge \Delta < a \\ \text{(iv)} \; \frac{(a-\Delta)(3a+\Delta+2k)-2(5a-\Delta+k)t+7t^2}{2(a-\Delta)(3a+\Delta+2k)-4(4a-\Delta+k)t+10t^2} \quad \text{if } \frac{(3a+\Delta-4t)(t-a+\Delta)}{2(a-\Delta)} \leq k \wedge \Delta \geq a \end{cases}$$

and if $\Delta \geq t + a$:

$$\hat{\delta}_{OC2} = \begin{cases} \text{(i)} \; \frac{(2k-3(t-a)+\Delta)(t^2-a^2+\Delta^2)}{5a^3+\Delta^3+\Delta^2(2k-3t)+9\Delta t^2+t^2(2k+5t)-a^2(9\Delta+2k+5t)+a(3\Delta^2-5t^2)} \\ \qquad \text{if } k < \frac{3}{2}(t-a) - \frac{\Delta}{2} + \frac{2(t^2-a^2)}{\Delta} \\ \text{(ii)} \; \frac{25a^2+9\Delta^2+10\Delta(2k-3t)+6a(5\Delta+2k-3t)+(2k-7t)(2k+t}{a^2+17\Delta^2+36\Delta k+4k^2-46\Delta t+4kt-31t^2+a(46\Delta-4k+30t)} \\ \qquad \text{if } \frac{3}{2}(t-a) - \frac{\Delta}{2} + \frac{2(t^2-a^2)}{\Delta} \leq k < \frac{7}{2}(t-a) - \frac{\Delta}{2} \\ \text{(iii)} \; \frac{(a-\Delta)(3a+\Delta+2k)-2(5a-\Delta+k)t+7t^2}{2(a-\Delta)(3a+\Delta+2k)-4(4a-\Delta+k)t+10t^2)} \quad \text{if } \frac{7}{2}(t-a) - \frac{\Delta}{2} \leq k \end{cases}$$

A.4 Critical Discount Factor for Collusion only on Side 1

Note that collusion on side 1 alone is only profitable if $\Delta \leq t - a$. Then,

$$\hat{\delta}_{OC1} = \begin{cases} \text{(i)} \ \frac{(2k-3(t-a)-\Delta)(t^2-a^2+\Delta^2)}{(a+\Delta)(5a^2+4a\Delta-\Delta^2-2ak+2\Delta k)-(5a^2+3\Delta^2)t+(-5a-9\Delta+2k)t^2+5t^3} \\ \quad \text{if } k < \frac{7}{2}t - \frac{3}{2}a + \frac{\Delta}{2} - \frac{4a\Delta t}{t^2-a^2+a\Delta} \\ \text{(ii)} \ \frac{\begin{bmatrix} (a+\Delta)^2(-3a+\Delta-2k)^2 - 2(a+\Delta)(7a+3\Delta)(3a-\Delta+2k)t \\ +(7a+3\Delta)^2t^2 + 8(3a-\Delta+2k)t^3 - 40t^4 \end{bmatrix}}{\begin{bmatrix} (a+\Delta)^2(-3a+\Delta-2k)^2 - 2(a+\Delta)(7a+3\Delta)(3a-\Delta+2k)t \\ +(25a^2+26a\Delta+17\Delta^2-16(a+\Delta)k)t^2 + 8(9a+\Delta+4k)t^3 - 64t^4 \end{bmatrix}} \\ \quad \text{if } \frac{7}{2}t - \frac{3}{2}a + \frac{\Delta}{2} - \frac{4a\Delta t}{t^2-a^2+a\Delta} \leq k < \frac{4t^2-3a^2-2a\Delta+\Delta^2+7at+3\Delta t}{2(a+\Delta)} \\ \text{(iii)} \ \frac{(a+\Delta)(3a-\Delta+2k)+2(k-2(a+\Delta))t-7t^2}{2(a-\Delta+2k-5t)t} \qquad \text{if } \frac{4t^2-3a^2-2a\Delta+\Delta^2+7at+3\Delta t}{2(a+\Delta)} \leq k \end{cases}$$

A.5 Collusion under Multi-Homing in case that $k \geq 2t - a$

Under this assumption, all consumers on side 1 will multi-home, i.e. $n_1^A = n_1^B = 1$ under Nash competition and price coordination. This will simplify demand on side 2 to be equal to $n_2^i = 1/2 + (p_2^j - p_2^i)/(2t)$ with $i \in \{1, 2\}, i \neq j$. Hence, Nash prices and profits result from extracting all utility from the consumer located at the opposite end of the Hotelling line on side 1, i.e. setting $p_1^i = k + a_1 n_2^i - t$, and maximizing profits $\pi_{MH}^i = 1p_1^i + n_2^i p_2^i$ with respect to p_2^i. Thus,

$$p_1^N = k + \frac{a_1}{2} - t \ ; \ p_2^N = t - a_1 \ ; \ \pi_{MH}^N = k - \frac{t}{2} \tag{A.1}$$

Hence, Nash profits do not depend on network effects if $k \geq 2t - a$, because all the additional profits that can be earned on the multi-homing side are given away to attract single-homing consumers on side 2.

Turning to price coordination, platforms will agree to extract all surplus from the indifferent side-2 consumer located at 1/2, while prices on side 1 cannot increase any further. In detail, collusive prices and profits are given as follows:

$$p_1^C = k + \frac{a_1}{2} - t \ ; \ p_2^C = k + a_2 - \frac{t}{2} \ ; \ \pi_{MH}^C = \frac{3}{2}k + \frac{a_1 + a_2}{2} - \frac{5}{4}t \tag{A.2}$$

The collusive profit increases in both network effects a_1 and a_2, which is the result of pure price effects. Demand does not change when moving from Nash to collusion, but platforms are able to set higher prices on the single-homing side. The willingness-to-pay on side 2 increases with both network benefits. Thus, platforms can demand a higher collusive price p_2^C if a_1 and/or a_2 increase. It is therefore straightforward that the gain from collusion also increases in the network benefits given that Nash profits are independent of a_1, a_2.

Let me now turn to optimal defection and assume once again that $a_1 = a + \Delta$ and $a_2 = a - \Delta$. As already pointed out in section 2.4, defection is not as attractive if $k \geq 2t - a$ than it is if $n_1^N = n_1^C < 1$. This is due to the fact that deviation cannot increase demand on side 1 any further. In consequence, lowering prices on side-2 will not induce a positive feedback loop as it was the case in sections 2.3 and 2.4. Instead, it might even be the case that there is no profitable deviation from collusive prices at all, namely when $k < (3t^2 - a^2 + \Delta^2 - 4at)/(2t)$. Thus, if this condition holds true, prices and profits are identical to the ones given in (A.2) and the critical discount factor is equal to zero. If $k \geq (3t^2 - a^2 + \Delta^2 - 4at)/(2t)$, however, optimal defection actually increases deviation profits over π_{MH}^C. In detail, deviation prices are identical to the ones outlined in section 2.4 given that $n_1^D = 1$ or even both deviation demands are equal to one. Thus, deviation profits are as follows:

$$\pi_{MH}^D = \begin{cases} \frac{3}{2}k + a - \frac{5}{4}t & \text{if } k < \frac{3t^2 - a^2 + \Delta^2 - 4at}{2t} \\ \frac{2t(t^2 - a^2 + \Delta^2)(10k + 4a - 15t) + 4t^2(2k(k + 2t + 2a) - (k - 2\Delta)(k + 2\Delta)) - (t^4 - (a^2 - \Delta^2)^2)}{16t(2t^2 - a^2 + \Delta^2)} & \\ & \text{if } k_1^* \leq k < k_2^{**} \wedge k \geq \frac{3t^2 - a^2 + \Delta^2 - 4at}{2t} \\ 2(k + a - t) - \frac{t^2 - a^2 + \Delta^2}{2t} & \text{if } k_2^{**} < k \wedge k \geq \frac{3t^2 - a^2 + \Delta^2 - 4at}{2t} \end{cases}$$

A platform's gain from deviation $(\pi_{MH}^D - \pi_{MH}^C)$ is therefore given as:

$$\pi_{MH}^D - \pi_{MH}^C = \begin{cases} 0 & \text{if } k < \frac{3t^2-a^2+\Delta^2-4at}{2t} \\ \frac{(a^2-\Delta^2+4at+(2k-3t)t)^2}{16t(2t^2-a^2+\Delta^2)} & \text{if } k_1^* \leq k < k_2^{**} \wedge k \geq \frac{3t^2-a^2+\Delta^2-4at}{2t} \\ \frac{2a^2-2\Delta^2+4at+2kt-5t^2}{4t} & \text{if } k_2^{**} < k \wedge k \geq \frac{3t^2-a^2+\Delta^2-4at}{2t} \end{cases}$$

and one can easily show that it is increasing in a given assumption 2.3, $k \geq 2t - a$ and the above boundaries. Moreover, $(\pi_{MH}^D - \pi_{MH}^C)$ falls in Δ. Thus, the same results as in section 2.4 prevail. Let me finally state the critical discount factor above which collusion upon monopoly prices will be an equilibrium of the repeated game:

$$\hat{\delta}_{MH} = \begin{cases} 0 & \text{if } k < \frac{3t^2-a^2+\Delta^2-4at}{2t} \\ \frac{(a^2-\Delta^2+4at+(2k-3t)t)^2}{(a^2-\Delta^2)^2-4(a-\Delta)(a+\Delta)(2a+k)t+2(11a^2-3\Delta^2+8ak+2k^2)t^2+4(2a+k)t^3-15t^4} & \\ & \text{if } k_1^* \leq k < k_2^{**} \wedge k \geq \frac{3t^2-a^2+\Delta^2-4at}{2t} \\ 1 - \frac{(2k-3t+4a)t}{2(a^2-\Delta^2+4at+2(k-2t)t)} & \text{if } k_2^{**} < k \wedge k \geq \frac{3t^2-a^2+\Delta^2-4at}{2t} \end{cases}$$

Analyzing this discount factor with respect to a and Δ, I confirm the results of section 2.4, namely that collusion becomes harder when total network benefits a increase, while it becomes easier to sustain monopoly prices when network effects are more asymmetric.

Appendix B

Appendix of Chapter 3

B.1 Proof of Lemma 1

Consider the bargaining situation in period 2, in which all investments i_1, i_2, I_1, I_2 are fixed. Therefore, the total period 2 revenues to be distributed amount to $V(i_1, i_2, I_1, I_2)$, with D able to obtain $V(i_1, i_2, 0, 0)$ on its own.

Lemma 1: *The following is the unique subgame-perfect Nash equilibrium of sequential bargaining in period 2 under knowledge diffusion: Firm 1 demands* $B_1 = V(i_1, i_2, I_j, 0) - V(i_1, i_2, 0, 0)$*, firm 2 demands* $B_2 = V(i_1, i_2, I_1, I_2) - V(i_1, i_2, I_j, 0)$ *with* $j = 1$ *and* D *accepts both offers.*

Proof: Consider firm 2's situation given that D has accepted B_1. If it demands any $B_2' > B_2$, then D will reject the offer. For any $B_2'' < B_2$, raising the demand to B_2 increases firm 2's payoff. D is indifferent between accepting and rejecting B_2. By the same arguments, if D has rejected B_1 previously, then firm 2 demands $B_2^- = V(i_1, i_2, 0, I_2) - V(i_1, i_2, 0, 0)$, instead, and D accepts. It is helpful to note that B_2 does not depend on B_1, apart from the success of the first stage of negotiation.

Now consider firm 1's offer to D. If it offers B_1, D anticipates B_2 and is indifferent between accepting and rejecting the offer. For $B_1' < B_1$, firm 1 could increase its profits by demanding B_1. If firm 1 demands $B_1'' > B_1$, then D faces the following problem. If it accepts, then firm 2 will demand B_2 and D makes overall profits below $V(i_1, i_2, 0, 0)$ – it would be better off rejecting B_1 and accepting B_2^-, or rejecting all offers. The equilibrium outlined in the Lemma is unique.

B.2 Proof of Lemma 2

Consider the bargaining situation in period 1, in which the investments i_1, i_2 are fixed. The period 1 revenues to be distributed amount to $v(i_1, i_2)$, with D able to obtain 0 on its own.

Lemma 2: *The following is the unique subgame-perfect Nash equilibrium of sequential bargaining in period 1: firm 1 demands $b_1 = 0$, firm 2 demands $b_2 = v(i_1, i_2)$ and D accepts both offers.*

Proof: If D has rejected b_1, then there is no surplus to be shared. Consider firm 2's offer, given that D has accepted b_1. If D rejects firm 2's offer, its period 1 surplus is $-b_1$. Therefore firm 2 optimally makes an offer such that D is indifferent between accepting and rejecting, by demanding $v(i_1, i_2)$. Anticipating this, the only non-negative offer by firm 1 that D is willing to accept is 0, as it can achieve 0 on its own by rejecting all offers.

B.3 Stochastic Order Bargaining under Permanent IP Rights

Under permanent IP rights, the upstream firm who makes the second offer is the residual claimant with the outside option of the other parties equal to 0, so that there is a complete hold-up. With the entire surplus of the respective period being allocated according to the shares $(p_1^t, 1-p_1^t)$, the outcome resembles Nash bargaining over the generated surplus between the upstream firms with the bargaining power parameter p. We can therefore derive the expected profit functions:

$$\pi_1 = (1-p_1^1)v(i_1,i_2) - c(i_1) + \beta[(1-p_1^2)V(i_1,i_2,I_1,I_2) - c(I_1)] \quad \text{(B.1)}$$

$$\pi_2 = (p_1^1)v(i_1,i_2) - c(i_2) + \beta[p_1^2 V(i_1,i_2,I_1,I_2) - c(I_2)] \quad \text{(B.2)}$$

$$\pi_D = 0 \quad \text{(B.3)}$$

This gives us the following first-order conditions for second-period investments:

$$(1-p_1^2)\frac{\partial V(i_1,i_2,I_1,I_2)}{\partial I_1} = c \quad \text{(B.4)}$$

$$p_1^2\frac{\partial V(i_1,i_2,I_1,I_2)}{\partial I_2} = c \quad \text{(B.5)}$$

For the first period investments, the following must hold:

$$(1-p_1^1)\frac{\partial v(i_1,i_2)}{\partial i_1} + \beta[(1-p_1^2)(\frac{\partial V(i_1,i_2,I_1^*(i_1,i_2),I_2^*(i_1,i_2))}{\partial i_1} + \frac{\partial V(i_1,i_2,I_1^*(i_1,i_2),I_2^*(i_1,i_2))}{\partial I_2}\frac{\partial I_2}{\partial i_1})] = c \quad \text{(B.6)}$$

$$p_1^1\frac{\partial v(i_1,i_2)}{\partial i_1} + \beta[p_1^2(\frac{\partial V(i_1,i_2,I_1^*(i_1,i_2),I_2^*(i_1,i_2))}{\partial i_2} + \frac{\partial V(i_1,i_2,I_1^*(i_1,i_2),I_2^*(i_1,i_2))}{\partial I_1}\frac{\partial I_1}{\partial i_2})] = c \quad \text{(B.7)}$$

Appendix C

Appendix of Chapter 4

C.1 Generating the Data Set

Data Sources We assembled the final data from various data sources that were merged after a thorough data cleaning process. We use:

1. Electronic cashier system data from a German restaurant from 10/24/2001 until 05/31/2009.

2. PDP-files of all menus handed out during the above period.

3. Official data on local weather conditions, unemployment rates, price indices and data from the CESIfo Group Munich on the "Ifo Business Climate Germany".

The first data source supplies us with text files of all bills issued during the period. Those bills in raw format include information on the number and prices of dishes and drinks ordered, the total billing amount, exact date and time, table number and waiter name. The second data set, namely the original menu copies, provides information on which dishes and drinks were offered during a specific menu period. Moreover, the original menu files allow us to manually construct variables that measure how the information is displayed to the customer. Third, we use supplementary official data to control for outside factors that might influence consumers' choice of products. In particular, we merge detailed regional weather data, such as daily information on temperature, height of precipitation and sunshine duration.[1] Furthermore, we attach measures for macroeconomic conditions, such as the unemployment rate on rural district level and an index measuring the business climate on national level to the final data set.[2] Last, we collect information about public holidays and school vacations.

Sample Selection To guarantee a clean analysis and proper identification, we drop several observations. First, we exclude all observations that were recorded before the introduction of the EURO. This ensures that there are no structural breaks in the data and that all changes in choice probabilities relate to changes in menu and not currency. Second, we drop bills of waiters for whom the data contain very few observations to make sure that our results are representative of a typical restaurant choice situation and not flawed by inexperienced waiter types. Third, we exclude all price-dish combinations that were chosen less than 10 times over the sample period. This includes special holiday offers and one-time price reductions, amounting to 9% of the overall sample. For the same reason we also exclude special group menus, wedding or all-you-can-eat buffets. Last, we exclude all items where prices were obviously misreported, i.e. where prices were different from any price recorded in the menu or where the price of an item is completey missing.

Menu Choice Variables Information about consumer choices and the corresponding menu design lies at the heart of this study. First, we manually group all chosen dishes as reported on their corresponding bills into the corresponding menu categories: starters, main dishes, children's dishes, desserts and drinks. Second, we sort all main dishes into the sub-categories defined in the menus. Sub-categories are traditional dishes, fish, venison, steaks, poultry and vegetarian dishes. Third, we account for the fact that 4 dishes were listed twice in the same menu (most of them both as a starter and as a main dish) by creating respective dummies. Last, we

[1] Source: Deutscher Wetterdienst (*German weather forecast services*), `www.dwd.de`.

[2] Regional unemployment rates have been obtained from *Bundesagentur für Arbeit: Arbeitslose nach Kreisen*, `http://statistik.arbeitsagentur.de`. The business climate index can be obtained from *CESifo GmbH: ifo Geschäftsklima für die Gewerbliche Wirtschaft*, `http://www.cesifo-group.de`.

generate indicator variables for dishes that were sold at a lower price as an elderly person's portion.

Price and Position Indicators Position indicators provide information about how the menu design influences choice. First and most importantly, we generate variables that contain information on the position of a certain dish within its category, where we alternatively assign low or high numbers to the first or to the last positions. Second, to account for possible nonlinearities, we include indicator variables for the first and last position in each group. Third, we generate indicator variables for whether a dish was recently added to the menu as a new item. Besides, we generate variables for whether an item was the first listed on a respective's menu page and whether it was highlighted "specialty of the house" or "home-made". Last, we include variables on how many items have been listed within a certain menu category.

In the data, an item's price and position in the menu are strongly correlated, such that price effects have to be carefully disentangled from position effects. Hence, next to the price information obtained from the restaurant bills, we also generate a second order price variable to account for nonlinear price effects. Moreover, we generate an indicator for dishes that cost more than 10 Euros, which might function as a psychological boundary.

Control Variables To make sure that choice situations are comparable over time, we control for the external setting in which the choice situations took place. First, we generate table fixed effects, indicating whether a table is located outside, on the ground floor level or on the first floor of the restaurant. We drop all bills for which the table number is missing. Second, we generate an indicator variable that equals one if a table was shared by more than one party, identified by whether two bills were issued on the same table at overlapping time periods.[3] Third, we include information on the overall business climate prevailing on a particular day. Fourth, we generated indicator variables for weekend days, public holidays, school vacation

[3]We also generated several variables that provide more detailed waiter information, such as a waiter sex, tenure or experience. These variables, however, proved to be largely unimportant for dish choice and were thus excluded from the analysis.

times, as well as variables that capture the day of the week and time of the day. Last, we coded a variable for the number of main dishes per bill, as a proxy of the number of customers sharing the same table.

C.2 Data Summary Tables

Table C.1: Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.
price_item	10.05	2.63	3.5	17.8
sum_category	4.92	1.44	2	8
rel_position	0.58	0.27	0.13	1
first_pos	0.23	0.42	0	1
last_pos	0.22	0.42	0	1
mid_pos	0.38	0.49	0	1
adav_median	1.16	0.85	0	3.5
smaller_median	0.53	0.5	0	1
bigger_median	0.6	0.49	0	1
evening	0.28	0.45	0	1
full_emp	0.71	0.46	0	1
trainee	0.19	0.39	0	1
temp_emp	0.11	0.31	0	1
ground	0.22	0.41	0	1
first	0.49	0.5	0	1
outside	0.29	0.46	0	1
double_sit	0.01	0.09	0	1
tm	11.56	7.11	-12.9	29.1
so	5.88	4.41	0	15.9
rr	1.26	3.51	0	46.4
unemp_rate	16.35	2.55	11.3	22.8
business climate	96.69	7.04	82.3	108.9
no_maindish	3.62	2.24	1	15
no_drink	5.16	4.35	0	63
hr_day	15.08	3.13	10	23
weekend	0.57	0.49	0	1
evening	0.28	0.45	0	1
public holiday	0.05	0.22	0	1
trad	0.28	0.45	0	1
fish	0.15	0.36	0	1
venison	0.16	0.36	0	1
steak	0.22	0.41	0	1
poultry	0.08	0.28	0	1
veggie	0.11	0.31	0	1
bison	0.2	0.4	0	1
asparagus	0.18	0.38	0	1
canarian	0.17	0.38	0	1
hiking	0.29	0.45	0	1
N		88113		

Table C.2: Summary statistics by categories

	traditional		**fish**		**venison**		**steak**		**poultry**		**vegetarian**	
	Mean	StdDev	Mean	StdDev	Mean	StdDev	Mean	StdDev	Mean	StdDev	Mean	StdDev
price item	8.85	2.01	11.07	1.64	12.95	1.82	10.75	1.58	9.51	2.34	6.36	2.00
sum category	6.71	1.06	4.46	0.70	4.10	0.30	4.45	1.05	3.43	0.54	4.23	0.84
rel position	0.61	0.29	0.59	0.26	0.59	0.26	0.56	0.25	0.60	0.25	0.52	0.29
first pos	0.14	0.35	0.19	0.40	0.25	0.43	0.23	0.42	0.32	0.47	0.42	0.49
last pos	0.31	0.46	0.18	0.38	0.16	0.37	0.19	0.39	0.26	0.44	0.17	0.38
mid pos	0.25	0.43	0.37	0.48	0.54	0.50	0.44	0.50	0.50	0.50	0.28	0.45
adav median	1.69	1.08	0.99	0.62	0.94	0.52	0.94	0.68	0.71	0.56	1.16	0.67
smaller median	0.63	0.48	0.51	0.50	0.46	0.50	0.50	0.50	0.56	0.50	0.41	0.49
bigger median	0.49	0.50	0.62	0.48	0.56	0.50	0.67	0.47	0.73	0.44	0.70	0.46
no maindish	3.32	2.04	3.76	2.29	3.95	2.32	3.72	2.32	3.77	2.35	3.41	2.21
outside	0.34	0.47	0.27	0.45	0.21	0.41	0.29	0.45	0.18	0.38	0.41	0.49
business climate	97.01	6.85	96.82	7.04	96.30	6.86	96.93	7.16	94.35	6.75	97.65	7.39
evening	0.17	0.37	0.29	0.45	0.25	0.44	0.34	0.48	0.32	0.47	0.39	0.49
public holiday	0.05	0.22	0.05	0.22	0.06	0.23	0.05	0.22	0.05	0.23	0.05	0.21
weekend	0.56	0.50	0.56	0.50	0.62	0.48	0.55	0.50	0.61	0.49	0.53	0.50

Table C.3: Observations for each menu period by dish category

	dish category						
	traditional	fish	venison	steak	poultry	vegetarian	Total
menu 1	433	243	320	361	289	160	1,806
menu 2	1,975	1,303	1,424	1,538	743	647	7,630
menu 3	513	339	407	368	436	120	2,183
menu 4	816	420	505	622	451	179	2,993
menu 5	2,024	1,104	1,171	1,376	901	659	7,235
menu 6	879	557	600	729	526	239	3,530
menu 7	2,702	1,330	1,396	1,639	679	679	8,425
menu 8	1,110	615	645	775	492	279	3,916
menu 9	2,026	868	946	1,247	463	686	6,236
menu 10	912	410	548	530	192	246	2,838
menu 11	619	395	447	447	310	200	2,418
menu 12	3,219	1,533	1,427	2,434	0	1,342	9,955
menu 13	849	576	604	688	693	346	3,756
menu 14	1,948	1,088	749	1,476	0	1,003	6,264
menu 15	723	355	353	639	0	326	2,396
menu 16	562	410	420	683	539	313	2,927
menu 17	1,509	924	666	1,890	0	856	5,845
menu 18	743	357	389	659	0	317	2,465
menu 19	380	360	342	527	436	261	2,306
menu 20	263	139	157	317	124	158	1,158
menu 21	484	242	247	457	162	239	1,831
Total	24,689	13,568	13,763	19,402	7,436	9,255	88,113

Table C.4: Number of available choices for each menu by dish category

	dish category						
	traditional	fish	venison	steak	poultry	vegetarian	Total
menu 1	7	4	4	3	3	3	24
menu 2	7	4	4	3	3	3	24
menu 3	7	4	4	3	4	3	25
menu 4	7	4	4	3	4	3	25
menu 5	7	4	4	3	3	4	25
menu 6	7	4	4	4	3	3	25
menu 7	8	4	4	4	3	3	26
menu 8	8	4	4	4	4	4	28
menu 9	8	4	4	4	3	4	27
menu 10	8	5	4	4	3	5	29
menu 11	8	5	4	4	4	4	29
menu 12	6	5	4	5	0	4	24
menu 13	6	5	5	4	4	4	28
menu 14	6	5	4	5	0	5	25
menu 15	6	5	4	6	0	5	26
menu 16	5	5	4	5	4	5	28
menu 17	5	5	4	6	0	5	25
menu 18	5	5	4	6	0	5	25
menu 19	5	6	5	5	4	6	31
menu 20	5	6	5	6	4	6	32
menu 21	5	5	5	6	4	5	30
Total choices	14	23	9	6	8	16	76

Table C.5: Variable names and explanations

variable name	**variable label**
price_item	price of item (per unit)
sum_category	total number of dishes in the respective dish category
rel_position	relative position within dish category
first_pos	first position within dish category
first_maindish	cross term measuring interaction of first_pos*no_maindish
last_pos	last position within dish category
last_maindish	cross term measuring interaction of last_pos*no_maindish
mid_pos	median position within dish category
mid_maindish	cross term measuring interaction of mid_pos*no_maindish
adav_median	absolute value of the absolute deviation from median position
smaller_median	dummy indicating if the position is lower than median position
bigger_median	dummy indicating if the position is higher than median position
evening	dummy indicating that the dish was ordered past 5 pm.
full_emp	waiters with full employment
trainee	waiters that are trainees
temp_emp	waiters with temporary/seasonal employment
ground	bills on ground floor
first	tables on first floor
outside	outside table
double_sit	tables with multiple parties
tm	average temperature 2m above ground
so	daily hours of sunshine
rr	daily precipitation height
unemp_rate	monthly unemployment rate at community level
business climate	monthly business climate indicator (national level)
no_maindish	total number of main dishes on the bill - group size proxy 1
no_drink	total number of drinks on the bill - group size proxy 2
hr_day	hour of the day when the bill was issued
weekend	dummy indicating that the dish was ordered on a weekend day
evening	dummy indicating that the dish was ordered past 5 pm.
public holiday	dummy indicating that the dish was ordered on a public holiday
trad	traditional dishes category
fish	fish dishes category
venison	venison dishes category
steak	steak dishes category
poultry	poultry dishes category
veggie	vegetarian dishes category
bison	dummy indicating that Bisonbraten was listed as a specialty on a separate page
asparagus	dummy indicating the presence of an extra asparagus menu
canarian	dummy indicating the presence of an extra Canarian Islands specialties menu
hiking	dummy indicating the presence of an extra menu with dishes for hikers

C.3 Regression Results

Table C.6: Compromise effects regressions (without controls)

	case 1	case 2	case 3	case 4	case 5	case 6	case 7	case 8	case 9
menu_change	0.2187	0.3726	0.3259	0.1004	**0.1312****	**1.3101*****	**.5624****	0.1781	0.4474
	(0.339)	(0.423)	(0.102)	(0.774)	(0.019)	(0.000)	(0.013)	(0.504)	(0.148)
marginal effect at the mean	0.0210	0.0455	**0.0496***	-0.0170	**0.0471*****	**0.2269*****	**0.0811****	**0.0435***	0.0543
	(0.197)	(0.321)	(0.077)	(0.666)	(0.000)	(0.000)	(0.019)	(0.073)	(0.137)
_c*	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
outside	No	No	No	No	No	No	No	No	No
business climate	No	No	No	No	No	No	No	No	No
public holiday	No	No	No	No	No	No	No	No	No
weekend	No	No	No	No	No	No	No	No	No
evening	No	No	No	No	No	No	No	No	No
no_maindish	No	No	No	No	No	No	No	No	No
Prob > chi2	0.2121	0.7506	0.0052	0.8124	0.0000	0.0000	0.0413	0.000	0.2080
Log Likelihood	-1841.7292	-590.96752	-1329.9726	-703.19099	-9801.8057	-478.44243	-657.1973	-1303.3907	-655.76913
N obs	5,990	1,815	3,936	2,100	27,363	1,434	1,902	4,368	2,270
N cases	1,198	363	984	525	9,121	478	634	1,092	454

The case-specific coefficient 'menu_change' as well as the 'marginal effect at the mean' both refer to the alternative that was exposed to the compromise effect in the specific case (see alternatives in bold in table 4.3).
Note: p-values in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.7: Compromise effects regressions (with controls)

	case 1	case 2	case 3	case 4	case 5	case 6	case 7	case 8	case 9
menu_change	0.2063	0.0323	**0.9035****	-0.7562	0.1201	**1.8676*****	**0.7235***	**0.9673****	**1.6863****
	(0.412)	(0.973)	(0.030)	(0.304)	(0.178)	(0.000)	(0.063)	(0.026)	(0.044)
marginal effect at the mean	0.0154	-0.0695	0.0162	-0.0927	**0.0313****	**0.2767*****	0.0473	0.0823	0.0716
	(0.378)	(0.453)	(0.792)	(0.188)	(0.038)	(0.000)	(0.433)	(0.441)	(0.992)
_c*	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
outside	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
business climate	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
public holiday	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
weekend	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
evening	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
no_maindish	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Prob > chi2	0.0002	-	0.0000	0.0003	0.0000	0.0001	0.0003	0.0000	0.0001
Log Likelihood	-1812.9257	-473.48776	-1289.9225	-677.65195	-9603.446	-468.94195	-639.46458	-1232.3855	-618.93288
N obs	5,990	1,815	3,936	2,100	27,363	1,434	1,902	4,368	2,270
N cases	1,198	363	984	525	9,121	478	634	1,092	454

The case-specific coefficient 'menu_change' as well as the 'marginal effect at the mean' both refer to the alternative that was exposed to the compromise effect in the specific case (see alternatives in bold in table 4.3).
Note: p-values in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.8: Conditional logit category 1 (traditional dishes)

	(1)	(2)	(3)	(4)	(5)	(6)
choice						
lprice_item	-0.241***	-0.434***	-0.0618*	-0.0724*	-0.0655*	-0.0844**
	(0.00)	(0.00)	(0.08)	(0.06)	(0.06)	(0.02)
first_pos	-0.0807*	-0.0949*				
	(0.08)	(0.06)				
last_pos	0.795***	0.981***				
	(0.00)	(0.00)				
mid_pos			0.0370	0.0421		
			(0.14)	(0.11)		
adav_median					0.0247	-0.0103
					(0.14)	(0.59)
_c*	Yes	Yes	Yes	Yes	Yes	Yes
business climate	No	Yes	No	Yes	No	Yes
weekend	No	Yes	No	Yes	No	Yes
evening	No	Yes	No	Yes	No	Yes
no_maindish	No	Yes	No	Yes	No	Yes
N	165571	165571	165571	165571	165571	165571
Log lik.	-44441.1	-43837.9	-44699.8	-44142.7	-44699.8	-44143.8
AIC	88914.2	87811.9	89429.5	88419.4	89429.6	88421.6
BIC	89074.5	88493.0	89579.8	89090.5	89579.8	89092.7

p-values in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Own data, 2002-2009.

Table C.9: Conditional logit category 2 (fish)

	(1)	(2)	(3)	(4)	(5)	(6)
choice						
lprice_item	-0.119***	-0.0508	-0.225***	-0.186***	-0.216***	-0.191***
	(0.00)	(0.30)	(0.00)	(0.00)	(0.00)	(0.00)
first_pos	-0.0781*	-0.114**				
	(0.07)	(0.03)				
last_pos	-0.426***	-0.485***				
	(0.00)	(0.00)				
mid_pos			0.285***	0.304***		
			(0.00)	(0.00)		
adav_median					-0.240***	-0.257***
					(0.00)	(0.00)
_c*	Yes	Yes	Yes	Yes	Yes	Yes
business climate	No	Yes	No	Yes	No	Yes
weekend	No	Yes	No	Yes	No	Yes
evening	No	Yes	No	Yes	No	Yes
no_maindish	No	Yes	No	Yes	No	Yes
N	61560	61560	61560	61560	61560	61560
Log lik.	-19208.2	-18843.7	-19222.6	-18862.3	-19214.7	-18853.4
AIC	38464.4	37903.5	38491.2	37938.6	38475.4	37920.8
BIC	38681.1	38878.5	38698.9	38904.5	38683.0	38886.8

p-values in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Own data, 2002-2009.

Table C.10: Conditional logit category 3 (venison)

	(1)	(2)	(3)	(4)	(5)	(6)
choice						
lprice_item	-0.460***	-0.178***	-0.326***	-0.383***	-0.323***	-0.381***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
first_pos	0.516***	0.386***				
	(0.00)	(0.00)				
last_pos	-1.818***	-2.066***				
	(0.00)	(0.00)				
mid_pos			-0.0321	0.0136		
			(0.24)	(0.71)		
adav_median					0.00628	-0.0148
					(0.82)	(0.69)
_c*	Yes	Yes	Yes	Yes	Yes	Yes
business climate	No	Yes	No	Yes	No	Yes
weekend	No	Yes	No	Yes	No	Yes
evening	No	Yes	No	Yes	No	Yes
no_maindish	No	Yes	No	Yes	No	Yes
N	56402	56402	56402	56402	56402	56402
Log lik.	-17986.2	-17668.1	-18829.9	-18605.7	-18830.5	-18605.7
AIC	35994.5	35422.2	37679.7	37295.4	37681.0	37295.4
BIC	36092.8	35806.6	37769.1	37670.9	37770.4	37670.9

p-values in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Own data, 2002-2009.

Table C.11: Conditional logit category 4 (steaks)

	(1)	(2)	(3)	(4)	(5)	(6)
choice						
lprice_item	-0.376***	-0.262***	-0.357***	-0.243***	-0.409***	-0.321***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
first_pos	-1.931***					
	(0.00)					
last_pos	-0.0324					
	(0.40)					
mid_pos			0.115***	0.113***		
			(0.00)	(0.00)		
adav_median					-0.185***	-0.214***
					(0.00)	(0.00)
_c*	Yes	Yes	Yes	Yes	Yes	Yes
business climate	No	Yes	No	Yes	No	Yes
weekend	No	Yes	No	Yes	No	Yes
evening	No	Yes	No	Yes	No	Yes
no_maindish	No	Yes	No	Yes	No	Yes
N	86387	86387	86387	86387	86387	86387
Log lik.	-26999.6	-26654.4	-26991.5	-26647.7	-26989.7	-26644.9
AIC	54013.3	53360.8	53997.1	53349.3	53993.5	53343.9
BIC	54078.8	53604.4	54062.7	53602.2	54059.0	53596.8

p-values in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Own data, 2002-2009.

Table C.12: Conditional logit category 5 (poultry)

	(1)	(2)	(3)	(4)	(5)	(6)
choice						
lprice_item	-0.125*	-0.104	-0.105	0.0122	-0.105	0.0122
	(0.09)	(0.28)	(0.14)	(0.90)	(0.14)	(0.90)
first_pos	-0.822***	-0.800***				
	(0.00)	(0.00)				
last_pos	-0.813***	-0.779***				
	(0.00)	(0.00)				
mid_pos			0.160***	0.0999		
			(0.01)	(0.12)		
adav_median					-0.160***	-0.0999
					(0.01)	(0.12)
_c*	Yes	Yes	Yes	Yes	Yes	Yes
business climate	No	Yes	No	Yes	No	Yes
weekend	No	Yes	No	Yes	No	Yes
evening	No	Yes	No	Yes	No	Yes
no_maindish	No	Yes	No	Yes	No	Yes
N	24928	24928	24928	24928	24928	24928
Log lik.	-8081.3	-7983.8	-8173.5	-8065.3	-8173.5	-8065.3
AIC	16182.6	16043.6	16365.0	16204.5	16365.0	16204.5
BIC	16263.9	16352.3	16438.2	16505.1	16438.2	16505.1

p-values in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Own data, 2002-2009.

Table C.13: Conditional logit category 6 (vegetarian dishes)

	(1)	(2)	(3)	(4)	(5)	(6)
choice						
lprice_item	0.150*	0.396***	0.246***	0.459***	0.307***	0.537***
	(0.07)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
first_pos	-0.267***	-0.154*				
	(0.00)	(0.06)				
last_pos	-0.0181	-0.0141				
	(0.78)	(0.84)				
mid_pos			0.0944**	0.113**		
			(0.04)	(0.02)		
adav_median					0.0181	0.104*
					(0.67)	(0.05)
_c*	Yes	Yes	Yes	Yes	Yes	Yes
business climate	No	Yes	No	Yes	No	Yes
weekend	No	Yes	No	Yes	No	Yes
evening	No	Yes	No	Yes	No	Yes
no_maindish	No	Yes	No	Yes	No	Yes
N	39134	39134	39134	39134	39134	39134
Log lik.	-12028.2	-11703.1	-12035.5	-11702.3	-12037.7	-11703.0
AIC	24092.4	23562.1	24105.1	23558.7	24109.3	23559.9
BIC	24246.7	24231.0	24250.9	24218.9	24255.1	24220.2

p-values in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Own data, 2002-2009.

Table C.14: Mixed logit category 1

	(1)	(2)	(3)	(4)	(5)	(6)
Mean						
lprice_item	0.0985***	-0.613***	0.165***	-0.0666*	0.239***	-0.0682*
	(0.00)	(0.00)	(0.00)	(0.06)	(0.00)	(0.05)
first_pos	-0.241***	-2.262***				
	(0.00)	(0.00)				
last_pos	0.143***	1.161***				
	(0.00)	(0.00)				
mid_pos			0.0477	-0.0137		
			(0.54)	(0.75)		
adav_median					-0.245***	0.0315*
					(0.00)	(0.08)
SD						
lprice_item	0.433***	0.814***	0.523***	-0.0571	0.765***	-0.0200
	(0.00)	(0.00)	(0.00)	(0.37)	(0.00)	(0.66)
first_pos	-0.0787	2.705***				
	(0.72)	(0.00)				
last_pos	0.00677	0.0166				
	(0.96)	(0.89)				
mid_pos			1.161***	-0.499***		
			(0.00)	(0.00)		
adav_median					0.0810	0.218***
					(0.26)	(0.00)
_c*	No	Yes	No	Yes	No	Yes
no_maindish	No	Yes	No	Yes	No	Yes
N	165571	165571	165571	165571	165571	165571
Log lik.	-46153.1	-44004.3	-46054.7	-44341.4	-46006.7	-44341.1
AIC	92318.3	88072.6	92117.3	88742.7	92021.4	88742.1
BIC	92378.4	88393.2	92157.4	89043.3	92061.5	89042.6

p-values in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Own data, 2002-2009.

Table C.15: Mixed logit category 2

	(1)	(2)	(3)	(4)	(5)	(6)
Mean						
lprice_item	-0.0790***	-0.0678	-0.0398***	-0.252***	-0.0496***	-0.221***
	(0.00)	(0.25)	(0.00)	(0.00)	(0.00)	(0.00)
first_pos	-0.516***	-0.124**				
	(0.00)	(0.03)				
last_pos	-0.716***	-0.548***				
	(0.00)	(0.00)				
mid_pos			0.173***	0.259***		
			(0.00)	(0.00)		
adav_median					-0.267***	-0.248***
					(0.00)	(0.00)
SD						
lprice_item	0.000900	0.256**	-0.000300	-0.000210	-0.000335	-0.0963
	(0.95)	(0.03)	(0.98)	(1.00)	(0.98)	(0.44)
first_pos	0.0290	0.143				
	(0.86)	(0.70)				
last_pos	1.219***	0.315				
	(0.00)	(0.40)				
mid_pos			-0.253	0.507**		
			(0.27)	(0.05)		
adav_median					0.00228	0.199
					(0.95)	(0.24)
_c*	No	Yes	No	Yes	No	Yes
no_maindish	No	Yes	No	Yes	No	Yes
N	61560	61560	61560	61560	61560	61560
Log lik.	-20100.9	-19164.8	-20337.6	-19178.6	-20185.6	-19171.3
AIC	40213.8	38425.5	40683.3	38449.2	40379.3	38434.5
BIC	40268.0	38858.8	40719.4	38864.4	40415.4	38849.8

p-values in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Own data, 2002-2009.

Table C.16: Mixed logit category 3

	(1)	(2)	(3)	(4)	(5)	(6)
Mean						
lprice_item	-0.0513	-0.344***	-0.101***	-0.329***	-0.122***	-0.327***
	(0.15)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
first_pos	-1.612***	0.104**				
	(0.00)	(0.04)				
last_pos	-39.61***					
	(0.00)					
mid_pos			0.273***	-0.0295		
			(0.00)	(0.30)		
adav_median					-0.350***	0.00164
					(0.00)	(0.96)
SD						
lprice_item	-1.390***	0.0302	0.0847	0.0427	0.234***	0.0633
	(0.00)	(0.67)	(0.12)	(0.59)	(0.00)	(0.47)
first_pos	3.253***	-0.172				
	(0.00)	(0.67)				
last_pos	32.04***					
	(0.00)					
mid_pos			0.398	0.0737		
			(0.11)	(0.69)		
adav_median					-0.0177	-0.0152
					(0.85)	(0.89)
_c*	No	Yes	No	Yes	No	Yes
no_maindish	No	Yes	No	Yes	No	Yes
N	56402	56402	56402	56402	56402	56402
Log lik.	-18144.1	-18793.8	-19103.6	-18796.5	-19071.9	-18797.1
AIC	36300.3	37627.5	38215.3	37633.0	38151.8	37634.2
BIC	36353.9	37806.3	38251.1	37811.8	38187.5	37813.0

p-values in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Own data, 2002-2009.

Table C.17: Mixed logit category 4

	(1)	(2)	(3)	(4)	(5)	(6)
Mean						
lprice_item	-0.140***	-0.403***	-0.147***	-0.442***	-0.143***	-0.414***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
first_pos	-1.273***					
	(0.00)					
last_pos	-0.468***	-3.823***				
	(0.00)	(0.00)				
mid_pos			0.257***	0.194***		
			(0.00)	(0.00)		
adav_median					-0.238***	-0.213***
					(0.00)	(0.00)
SD						
lprice_item	0.00226	-0.131	0.00196	0.129	0.00285	0.130
	(0.93)	(0.25)	(0.91)	(0.14)	(0.89)	(0.22)
first_pos	2.215***					
	(0.00)					
last_pos	0.000464	7.513***				
	(1.00)	(0.00)				
mid_pos			-0.00278	1.240***		
			(0.97)	(0.00)		
adav_median					-0.00122	0.605***
					(0.97)	(0.00)
_c*	No	Yes	No	Yes	No	Yes
no_maindish	No	Yes	No	Yes	No	Yes
N	86387	86387	86387	86387	86387	86387
Log lik.	-27273.9	-26818.8	-27542.8	-26889.8	-27461.6	-26895.2
AIC	54559.7	53665.6	55093.5	53807.7	54931.2	53818.4
BIC	54615.9	53796.7	55131.0	53938.8	54968.6	53949.5

p-values in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Own data, 2002-2009.

Table C.18: Mixed logit category 5

	(1)	(2)	(3)	(4)	(5)	(6)
Mean						
lprice_item	-0.209***	-0.0908	-0.205***	-0.112	-0.207***	-0.112
	(0.00)	(0.21)	(0.00)	(0.12)	(0.00)	(0.12)
first_pos	-0.755***	0.0108				
	(0.00)	(0.93)				
last_pos	-152.3					
	(0.35)					
mid_pos			0.845***	0.158***		
			(0.00)	(0.01)		
adav_median					-0.860***	-0.159***
					(0.00)	(0.01)
SD						
lprice_item	0.00143	0.00745	0.519***	0.00800	0.519***	0.00789
	.	(0.87)	(0.00)	(0.87)	(0.00)	(0.87)
first_pos	0.00279	0.00651				
	(0.98)	(0.97)				
last_pos	155.5					
	(0.35)					
mid_pos			1.833***	-0.0460		
			(0.00)	(0.92)		
adav_median					2.033***	0.0454
					(0.00)	(0.92)
_c*	No	Yes	No	Yes	No	Yes
no_maindish	No	Yes	No	Yes	No	Yes
N	24928	24928	24928	24928	24928	24928
Log lik.	-8156.5	-8167.2	-8402.5	-8163.8	-8402.1	-8163.8
AIC	16322.9	16370.3	16813.0	16363.6	16812.3	16363.6
BIC	16363.6	16516.6	16845.5	16509.8	16844.8	16509.8

p-values in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Own data, 2002-2009.

Table C.19: Mixed logit category 6

	(1)	(2)	(3)	(4)	(5)	(6)
Mean						
lprice_item	-0.393***	0.120	-0.278***	0.196*	-0.277***	0.185*
	(0.00)	(0.33)	(0.00)	(0.06)	(0.00)	(0.08)
first_pos	-0.417***	-0.314**				
	(0.00)	(0.02)				
last_pos	0.0442	-0.100				
	(0.50)	(0.32)				
mid_pos			-0.0411	-0.0330		
			(0.29)	(0.63)		
adav_median					0.0176	0.0425
					(0.62)	(0.50)
SD						
lprice_item	0.505***	-0.0359	0.388***	-0.0167	0.385***	-0.0152
	(0.00)	(0.71)	(0.00)	(0.81)	(0.00)	(0.82)
first_pos	0.583	0.610				
	(0.23)	(0.23)				
last_pos	-0.0292	-0.0102				
	(0.91)	(0.98)				
mid_pos			-0.0280	-0.0122		
			(0.90)	(0.95)		
adav_median					0.0241	0.0188
					(0.81)	(0.86)
_c*	No	Yes	No	Yes	No	Yes
no_maindish	No	Yes	No	Yes	No	Yes
N	24509	24509	24509	24509	24509	24509
Log lik.	-7440.9	-7370.1	-7455.7	-7376.3	-7456.1	-7376.2
AIC	14893.7	14788.3	14919.3	14796.7	14920.2	14796.4
BIC	14942.4	14982.8	14951.8	14975.0	14952.6	14974.8

p-values in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Source: Own data, 2002-2009.

Table C.20: Conditional logit full choice set

	(1)	(2)	(3)	(4)	(5)	(6)
choice						
lprice_item	-0.204**	-0.208***	-0.205***	-0.209***	-0.209***	-0.213***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
first_pos	0.0612	0.0558				
	(0.53)	(0.57)				
last_pos	-0.200***	-0.198***				
	(0.01)	(0.01)				
mid_pos			0.0306	0.0287		
			(0.61)	(0.63)		
adav_median					-0.0338	-0.0326
					(0.45)	(0.47)
_c*	Yes	Yes	Yes	Yes	Yes	Yes
no_maindish	No	Yes	No	Yes	No	Yes
N	82100	82100	82100	82100	82100	82100
Log lik.	-9969.7	-9900.8	-9973.5	-9904.5	-9973.4	-9904.3
AIC	20093.5	20103.6	20099.0	20108.9	20098.7	20108.6
BIC	20810.8	21510.3	20807.0	21506.3	20806.7	21506.0

p-values in parentheses
Source: Own data, random draw, 2002-2009.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table C.21: Mixed logit full choice set

	(1)	(2)	(3)	(4)	(5)	(6)
Mean						
lprice_item	0.00230	-0.216***	-0.0261***	-0.210***	-0.0200***	-0.212***
	(0.77)	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)
first_pos	0.0743	0.0494				
	(0.18)	(0.64)				
last_pos	-0.568***	-0.447***				
	(0.00)	(0.00)				
mid_pos			-0.0409	0.0268		
			(0.81)	(0.66)		
adav_median					-0.0539**	-0.0343
					(0.04)	(0.45)
SD						
lprice_item	0.0892***	-0.0922	0.0542	-0.0133	0.0534	-0.0121
	(0.00)	(0.40)	(0.12)	(0.88)	(0.13)	(0.89)
first_pos	0.142	0.173				
	(0.75)	(0.73)				
last_pos	0.773***	1.124***				
	(0.00)	(0.00)				
mid_pos			-1.122	-0.116		
			(0.12)	(0.78)		
adav_median					0.443***	0.129
					(0.00)	(0.55)
_c*	No	Yes	No	Yes	No	Yes
group size proxy	No	Yes	No	Yes	No	Yes
N	82100	82100	82100	82100	82100	82100
Log lik.	-10268.2	-9896.3	-10300.8	-9904.4	-10296.8	-9904.2
AIC	20548.4	20100.7	20609.6	20112.8	20601.6	20112.4
BIC	20604.3	21535.3	20646.8	21528.8	20638.8	21528.4

p-values in parentheses
Source: Own data, random draw, 2002-2009.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Bibliography

AGHION, P., AND J. TIROLE (1994): "The management of innovation," *The Quarterly Journal of Economics*, 109(4), 1185–1209.

ARGENTESI, E., AND L. FILISTRUCCHI (2007): "Estimating market power in a two-sided market: The case of newspapers," *Journal of Applied Econometrics*, 22(7), 1247–1266.

ARIELY, D., AND J. LEVAV (2000): "Sequential choice in group settings: Taking the road less traveled and less enjoyed," *Journal of Consumer Research*, 27(3), 279–290.

ARMSTRONG, M. (2006): "Competition in Two-Sided Markets," *The RAND Journal of Economics*, 37(3), 668–691.

ARMSTRONG, M., AND J. WRIGHT (2007): "Two-sided Markets, Competitive Bottlenecks and Exclusive Contracts," *Economic Theory*, 32(2), 353–380.

BELLEFLAMME, P., AND M. PEITZ (2010): "Platform competition and seller investment incentives," *European Economic Review*, 54(8), 1059–1076.

BESSEN, J., AND E. MASKIN (2009): "Sequential Innovation, Patents, and Imitation," *The RAND Journal of Economics*, 40(4), 611–635.

BORDALO, P. (2010): "Choice-Set Effects and Salience Theory," *Working Paper*.

BUDZINSKI, O., AND K. WACKER (2007): "The Prohibition of the Proposed Springer-ProSiebenSat.1-Merger: How Much Economics in German Merger Control?," *Journal of Competition Law & Economics*, 3(2), 281–306.

Buehler, S., and A. Schmutzler (2008): "Intimidating competitors - Endogenous vertical integration and downstream investment in successive oligopoly," *International Journal of Industrial Organization*, 26(1), 247–265.

Bundeskartellamt (1999): "Beschluss - Aktenzeichen: B6-94/10," `http://www.bundeskartellamt.de/wDeutsch/download/pdf/Kartell/Kartell03/B6_49_99.pdf?navid=36`, last access: November 10, 2009.

——— (2009): "Die Wettbewerbsaufsicht des Bundeskartellamtes - Schwerpunkte 2007/2008," `http://www.bundeskartellamt.de/wDeutsch/download/pdf/Taetigkeitsbericht/TB_Kurzfassung_07-08-D.pdf`, last access: December 10, 2011.

bwGRiD (http://www.bw-grid.de/) (2007-2010): "Member of the German D-Grid initiative, funded by the Ministry of Education and Research (Bundesministerium für Bildung und Forschung) and the Ministry for Science, Research and Arts Baden-Wuerttemberg (Ministerium für Wissenschaft, Forschung und Kunst Baden-Württemberg)," Universities of Baden-Württemberg.

Caillaud, B., and B. Jullien (2003): "Chicken & Egg: Competition among Intermediation Service Providers," *The RAND Journal of Economics*, 34(2), 309–328.

Cameron, A., and P. Trivedi (2009): *Microeconomics using Stata.* College Station, Tex.: Stata Press books.

Chandra, A., and A. Collard-Wexler (2009): "Mergers in Two-Sided Markets: An Application to the Canadian Newspaper Industry," *Journal of Economics & Management Strategy*, 18(4), 1045–1070.

Chang, M.-H. (1991): "The effects of product differentiation on collusive pricing," *International Journal of Industrial Organization*, 9(3), 453–469.

Choi, J. (2010): "Tying in Two-Sided Markets with Multi-Homing," *The Journal of Industrial Economics*, 58(3), 607–626.

Coase, R. (1972): "Durability and Monopoly," *Journal of Law and Economics*, 15(1), 143–149.

de Clippel, G., and K. Eliaz (2010): "Reason-based Choice: A Bargaining Rationale for the Attraction and Compromise Effects," *Theoretical Economics*, forthcoming.

Dekel, E., and B. Lipman (2010): "How (Not) to Do Decision Theory," *Annual Review of Economics*, 2(1), 257–282.

DellaVigna, S. (2009): "Psychology and Economics: Evidence from the Field," *Journal of Economic Literature*, 47(2), 315–372.

Dewenter, R., J. Haucap, and T. Wenzel (2011): "Semi-collusion in media markets," *International Review of Law and Economics*, 31(2), 92–98.

Dhar, R., S. Nowlis, and S. Sherman (2000): "Trying hard or hardly trying: An analysis of context effects in choice," *Journal of Consumer Psychology*, 9(4), 189–200.

Doyle, J., D. O'Connor, G. Reynolds, and P. Bottomley (1999): "The robustness of the asymmetrically dominated effect: Buying frames, phantom alternatives, and in-store purchases," *Psychology and Marketing*, 16(3), 225–243.

European Commission (2002): "COMP/E-2/37.784 Fine Art Auction Houses," `http://ec.europa.eu/competition/antitrust/cases/decisions/37784/en.pdf`, last access: November 10, 2009.

Evans, D. S. (2003): "The Antitrust Economics of Two-Sided Markets," `http://ssrn.com/paper=363160`, Yale Journal of Regulation, last access: February 22, 2011.

Evans, D. S., and R. Schmalensee (2008): "Markets with Two-Sided Platforms," *Issues in Competition Law and Policy (ABA Section of Antitrust Law)*, 1(Chapter 28).

Fershtman, C., and S. Markovich (2010): "Patents, Imitation and Licensing in an Asymmetric Dynamic R&D Race," *International Journal of Industrial Organization*, 28(2), 113–126.

FISHMAN, A., AND R. ROB (2000): "Product innovation by a durable-good monopoly," *The RAND Journal of Economics*, 31(2), 237–252.

FRIEDMAN, J. W. (1971): "A Non-cooperative Equilibrium for Supergames," *The Review of Economic Studies*, 38(1), 1–12.

GAERTNER, W., AND Y. XU (1999): "On rationalizability of choice functions: A characterization of the median," *Social Choice and Welfare*, 16(4), 629–638.

GALLINI, N. T. (1992): "Patent Policy and Costly Imitation," *The RAND Journal of Economics*, 23(1), 52–63.

GIERL, H., AND K. STIEGELMAYR (2010): "Preis und Qualität als Dimensionen von Kompromissoptionen," *Zeitschrift für Betriebswirtschaft*, 80(5), 495–531.

GREEN, P., AND V. SRINIVASAN (1978): "Conjoint analysis in consumer research: issues and outlook," *Journal of Consumer Research*, pp. 103–123.

GROSSMAN, S., AND O. HART (1986): "The costs and benefits of ownership: A theory of vertical and lateral integration," *The Journal of Political Economy*, 94(4), 691–719.

HÄCKNER, J. (1995): "Endogenous product design in an infinitely repeated game," *International Journal of Industrial Organization*, 13(2), 277–299.

HANDELSBLATT (2002): "SZ und FR dürfen bei Anzeigen kooperieren," `http://www.handelsblatt.com/sz-und-fr-duerfen-bei-anzeigen-kooperieren/2181020.html`, last access: February 24, 2011.

HART, O., AND J. MOORE (1990): "Property Rights and the Nature of the Firm," *Journal of Political Economy*, 98(6), 1119–1158.

HUBER, J., J. PAYNE, AND C. PUTO (1982): "Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis," *Journal of Consumer Research*, 9(1), 90–98.

INDERST, R., AND N. MAZZAROTTO (2008): "Buyer power in distribution," in *ABA Antitrust Section Handbook, Issues in Competition Law and Policy*, vol. 3, pp. 1953–1978. ABA.

Inderst, R., and C. Wey (2003): "Bargaining, mergers, and technology choice in bilaterally oligopolistic industries," *RAND Journal of Economics*, 34(1), 1–19.

——— (2005): "How Strong Buyers Spur Upstream Innovation," *CEPR Discussion Papers*, http://www.cepr.org/pubs/dps/DP5365.asp.

——— (2007): "Buyer power and supplier incentives," *European Economic Review*, 51(3), 647–667.

——— (2011): "Countervailing power and dynamic efficiency," *Journal of the European Economic Association*, 9(4), 702–720.

Ishii, A. (2004): "Cooperative R&D between vertically related firms with spillovers," *International Journal of Industrial Organization*, 22(8-9), 1213–1235.

Josiam, B., and J. Hobson (1995): "Consumer choice in context: the decoy effect in travel and tourism," *Journal of Travel Research*, 34(1), 45–50.

Kahneman, D., J. Knetsch, and R. Thaler (1991): "Anomalies: The endowment effect, loss aversion, and status quo bias," *The Journal of Economic Perspectives*, 5(1), 193–206.

Kaiser, U., and J. Wright (2006): "Price structure in two-sided markets: Evidence from the magazine industry," *International Journal of Industrial Organization*, 24(1), 1–28.

Katz, M. L., and C. Shapiro (1985): "Network Externalities, Competition, and Compatibility," *American Economic Review*, 75(3), 424–440.

Kivetz, R., O. Netzer, and V. Srinivasan (2004a): "Alternative models for capturing the compromise effect," *Journal of Marketing Research*, 41(3), 237–257.

——— (2004b): "Extending compromise effect models to complex buying situations and other context effects," *Journal of Marketing Research*, 41(3), 262–268.

Lehmann, D., and Y. Pan (1994): "Context effects, new brand entry, and consideration sets," *Journal of Marketing Research*, 31(3), 364–374.

Levine, J., and R. Moreland (1998): "Small groups," *The Handbook of Social Psychology*, 2, 415–469.

Luce, R. (1959): *Individual choice behavior*. John Wiley.

——— (1977): "The choice axiom after twenty years," *Journal of Mathematical Psychology*, 15(3), 215–233.

Lülfesmann, C. (2001): "Team production, sequential investments, and stochastic payoffs," *Journal of Institutional and Theoretical Economics JITE*, 157(3), 430–442.

Mansfield, E. (1985): "How Rapidly Does New Industrial Technology Leak Out?," *The Journal of Industrial Economics*, 34, 217–223.

Masten, S. E. (1995): "Old school ties: financial aid coordination and the governance of higher education," *Journal of Economic Behavior & Organization*, 28(1), 23–47.

McFadden, D. (1973): "Conditional logit analysis of qualitative choice behavior," in *Frontiers in Econometrics*, ed. by P. Zarembka, pp. 105–142. Academic press.

——— (2001): "Economic choices," *American Economic Review*, 91(3), 351–378.

Merrilees, W. (1983): "Anatomy of a price leadership challenge: an evaluation of pricing strategies in the Australian newspaper industry," *The Journal of Industrial Economics*, 31(3), 291–311.

Nahm, J. (2004): "Durable-Goods Monopoly with Endogenous Innovation," *Journal of Economics & Management Strategy*, 13(2), 303–319.

Negrín, J. (2005): "The Regulation of Payment Cards: The Mexican Experience," *Review of Network Economics*, 4(4), 243–265.

Nöldeke, G., and K. Schmidt (1998): "Sequential investments and options to own," *RAND Journal of Economics*, 29(4), 633–653.

O'Muircheartaigh, C., J. A. Krosnick, and A. Helic (2000): "Middle Alternatives, Acquiescence, and the Quality of Questionnaire Data," Working Paper 0103, Harris School of Public Policy Studies, University of Chicago.

Peitz, M., and T. Valletti (2008): "Content and advertising in the media: Pay-tv versus free-to-air," *International Journal of Industrial Organization*, 26(4), 949–965.

Picard, R. (1995): "Free Press and Government: the ignored economic relationships of US newspapers," in *Media Structure and the State: Concepts, Issues, Messures.*, ed. by K. E. Gustafsson, pp. 143–148. Göteborg: Mass Media Unit, School of Economics and Commercial Law, Göteborg University.

Rochet, J.-C., and J. Tirole (2003): "Platform Competition in Two-Sided Markets," *Journal of the European Economic Association*, 1(4), 990–1029.

——— (2006): "Two-Sided Markets: A Progress Report," *RAND Journal of Economics*, 37(3), 645–667.

Rochet, J. C., and J. Tirole (2008): "Tying in two-sided markets and the honor all cards rule," *International Journal of Industrial Organization*, 26(6), 1333–1347.

Romeo, C., R. Pittman, and N. Familant (2003): "Do Newspaper JOAs Charge Monopoly Advertising Rates?," *Review of Industrial Organization*, 22(2), 121–138.

Rysman, M. (2009): "The Economics of Two-Sided Markets," *The Journal of Economic Perspectives*, 23(3), 125–143.

Salant, Y. (2011): "Procedural analysis of choice rules with applications to bounded rationality," *American Economic Review*, 101(2), 724–748.

Salant, Y., and A. Rubinstein (2008): "(A, f): Choice with Frames," *Review of Economic Studies*, 75(4), 1287.

Salop, S. C., and L. J. White (1991): "Policy Watch: Antitrust Goes to College," *Journal of Economic Perspectives*, 5(3), 193–202.

Sheng, S., A. Parker, and K. Nakamoto (2005): "Understanding the mechanism and determinants of compromise effects," *Psychology and Marketing*, 22(7), 591–609.

Simonson, I. (1989): "Choice based on reasons: The case of attraction and compromise effects," *Journal of Consumer Research*, 16(2), 158–174.

Simonson, I., and S. Nowlis (2000): "The role of explanations and need for uniqueness in consumer decision making: Unconventional choices based on reasons," *Journal of Consumer Research*, 27(1), 49–68.

Simonson, I., S. Nowlis, and K. Lemon (1993): "The effect of local consideration sets on global choice between lower price and higher quality," *Marketing Science*, 12(4), 357–377.

Simonson, I., and A. Tversky (1992): "Choice in context: Tradeoff contrast and extremeness aversion," *Journal of Marketing Research*, 29(3), 281–295.

Smirnov, V., and A. Wait (2004a): "Hold-up and sequential specific investments," *Rand Journal of Economics*, 35(2), 386–400.

——— (2004b): "Timing of investments, holdup and total welfare," *International Journal of Industrial Organization*, 22(3), 413–425.

Train, K. (2003): *Discrete choice methods with simulation*. Cambridge Univ Pr.

Tversky, A., and I. Simonson (1993): "Context-dependent preferences," *Management Science*, 39(10), 1179–1189.

Van Kranenburg, H. (2001): "Economic Effects of Consolidations of Publishers and Newspapers in The Netherlands," *The Journal of Media Economics*, 14(2), 61–76.

Van Kranenburg, H., F. Palm, and G. Pfann (2002): "Exit and survival in a concentrating industry: the case of daily newspapers in the Netherlands," *Review of Industrial Organization*, 21(3), 283–303.

Verband der Automobilindustrie (2011): "Meldung: Bundeskanzlerin Merkel eröffnet die weltweit wichtigste Mobilitätsmesse," `http://www.vda.de/de/meldungen/news/20110915-1.html`, last access: December 31, 2011.

Weiner, S., and J. Wright (2005): "Interchange fees in various countries: developments and determinants," *Review of Network Economics*, 4(4), 3.

Wernerfelt, B. (1995): "A rational reconstruction of the compromise effect: Using market data to infer utilities," *Journal of Consumer Research: An Interdisciplinary Quarterly*, 21(4), 627–633.

Wettbewerbskommission der Schweizerischen Eidgenossenschaft (2000a): "Jahresbericht 2000 der Wettbewerbskommission gemäss Art. 49 Abs. 2 KG," `http://www.weko.admin.ch/org/00143/index.html?lang=de&download=NHzLpZeg7t,lnp6I0NTU042l2Z6ln1acy4Zn4Z2qZpnO2Yuq2Z6gpJCDdHx,fWym162epYbg2c_JjKbNoKSn6A--`, last access: September 29, 2011.

——— (2000b): "Pressemitteilung 8.2.2000: Weko genehmigt einvernehmliche Regelung mit den drei Tessiner Tageszeitungen," `http://www.admin.ch/cp/d/389fe731.0@fwsrvg.bfi.admin.ch.html`, last access: September 29, 2011.

Weyl, E. (2010): "A Price Theory of Multi-sided Platforms," *American Economic Review*, 100(4), 1642–1672.

White, A., and E. G. Weyl (2010): "Imperfect Platform Competition: A General Framework," *NET Institute Working Paper No. 10-17*.

Wilbur, K. C. (2008): "A Two-Sided, Empirical Model of Television Advertising and Viewing Markets," *Marketing Science*, 27(3), 356–378.

Eidesstattliche Erklärung

Hiermit erkläre ich, die vorliegende Dissertation selbständig angefertigt und mich keiner anderen als der in ihr angegebenen Hilfsmittel bedient zu haben. Insbesondere sind sämtliche Zitate aus anderen Quellen als solche gekennzeichnet und mit Quellenangaben versehen.

Mannheim, 01.01.2012.

Isabel Ruhmer

Curriculum Vitae

1983	Born in Grimma, Germany
2002	Graduation from Secondary School, Johann-Gottfried-Seume Gymnasium, Grimma, Germany
2002–2005	Bachelor Studies in European Economic Studies, Otto-Friedrich-University Bamberg, Germany
2004–2005	Studies abroad, Université d'Angers, France
2005–2007	Master Studies in Economics, University of Duisburg-Essen, Germany
2007–2011	Graduate Studies in Economics, Center for Doctoral Studies in Economics (CDSE), University of Mannheim, Germany
2008–2011	Teaching Assistant for undergraduate and graduate level courses, University of Mannheim, Germany
2009	Graduate Studies in Economics, Toulouse School of Economics (visiting), University of Toulouse I, France