Two-Sided Markets in the Online World

An Empirical Analysis

D I S S E R T A T I O N

zur Erlangung des akademischen Grades

Dr. rer. pol.
im Fach Wirtschaftswissenschaften

eingereicht an der
Wirtschaftswissenschaftlichen Fakultät
Humboldt-Universität zu Berlin

von
Dipl.-Volksw. Thomas Hildebrand
geb. am 06.04.1981 in Berlin

Präsident der Humboldt-Universität zu Berlin:
Prof. Dr. Jan-Hendrik Olbertz

Dekan der Wirtschaftswissenschaftlichen Fakultät:
Prof. Oliver Günther, Ph.D.

Gutachter:
1. Prof. Lars-Hendrik Röller, Ph.D.
2. Prof. Jörg Rocholl, Ph.D.

eingereicht am: 14.03.2011
Tag der mündlichen Prüfung: 16.06.2011

To my family and Steffi.

Abstract

This dissertation investigates various aspects of two-sided markets – markets with at least two distinct user groups each of which exerts inter-group network effects on the other side – in the online world.

In the first paper, I examine the interactions between the traditional (offline) demand channels and the new (online) demand channels in the German magazines industry, focusing in particular on the link between offline and online advertising. I find that offline and online advertising are substitutes although not perfect ones. This explains the shift from offline to online advertising observed in recent years.

In the second paper, I develop a semi-structural approach to identify network effects on two-sided monopoly platforms without data on prices and quantities. A sufficient test for the existence of network effects is derived when only data on total revenue is available. If separate revenue data is available on the two sides, then the test is both necessary and sufficient.

The third paper investigates the certification mechanisms and incentives that enable lending markets to match demand and supply despite the absence of financial intermediaries with skin in the game. The institutional setting for this analysis is the online social lending market, in which potential lenders and borrowers interact directly without a financial intermediary but can create self-organized groups instead. A difference-in-difference approach is used to examine how the same groups behave before and after the exogenously imposed elimination of rewards for the leaders of these groups. Allowing group leader rewards is found to be detrimental for the market outcome. Group leaders become more careful in screening after the elimination of these rewards, and if their loan participation is high, i.e. when they have skin in the game and are thus severely hurt by a borrower default.

Zusammenfassung

Diese Dissertation besteht aus drei Aufsätzen, in denen verschiedene Aspekte von zweiseitigen Märkten untersucht werden. Dabei handelt es sich um Märkte mit zwei verschiedenen Nutzergruppen, von denen jede Netzwerkeffekte auf die jeweils andere Seite ausübt.

Im ersten Aufsatz werden die Wechselwirkungen zwischen den traditionellen (offline) Nachfragekanälen und den neuen (online) Nachfragekanälen in der deutschen Zeitschriftenindustrie analysiert. Dabei liegt der Fokus insbesondere auf den Effekten zwischen Offlinewerbung und Onlinewerbung. Das Ergebnis der Schätzung eines dafür entwickelten ökonomischen Modells ist, dass Offline- und Onlinewerbung moderate Substitute füreinander sind. Dies erklärt die Verlagerung von Offline- zu Onlinewerbung, die man in den vergangen Jahren beobachten konnte.

Im zweiten Aufsatz wird ein semistruktureller Ansatz zur Messung von Netzwerk-Effekten auf potentiell zweiseitigen Monopol-Plattformen entwickelt. Der Test ist hinreichend, wenn lediglich Daten zum Gesamtumsatz der Plattform zur Verfügung stehen. Sind getrennte Umsatzdaten für die beiden Seiten verfügbar, dann ist der Test sowohl notwendig als auch hinreichend.

Der dritte Aufsatz untersucht Mechanismen und Anreize, die die Koordination von Angebot und Nachfrage auf Kreditmärkten ermöglichen, in denen es keine Finanzintermediäre mit eigener finanzieller Beteiligung an den vergebenen Darlehen gibt. Dazu wird der Online-Direktkreditmarkt analysiert, in welchem an die Stelle von klassischen Finanzintermediären ein System von Gruppen tritt. Anhand eines Differenz-in-Differenzen-Ansatzes wird gezeigt, dass Entgelte für die Leiter dieser Gruppen zu adversen Anreizen führen können. Nach Abschaffung der Entgelte differenzieren die Leiter der Gruppen deutlich stärker bei der Auswahl derjenigen Kreditgesuche, die sie als investitionswürdig empfehlen. Gleiches ist zu beobachten, wenn die Leiter der Gruppen selbst zu einem großen Teil an den entsprechenden Darlehen beteiligt sind.

Acknowledgments

While writing this dissertation, I enjoyed the support of many people whom I would like to thank. First and foremost, I am particularly grateful to my supervisors Lars-Hendrik Röller and Jörg Rocholl. Both of them were always ready to provide profound and constructive feedback to my ideas and questions.

I would like to thank Özlem Bedre-Defolie, Francis Bidault, Michał Grajek, Paul Heidhues, and Manju Puri. My work very much benefited from their critical comments. In particular, the third paper of this dissertation is the result of joint work with Manju Puri and Jörg Rocholl.

Furthermore, I am also grateful to my fellows in the Berlin Doctoral Program in Economics and Management Science, in particular Dora Simroth, for fruitful and constructive discussions.

Financial support for this dissertation was provided by an Elsa-Neumann-Scholarship in the context of the NaFöG of the Land Berlin and is gratefully acknowledged.

Last but not least, I am thankful to my family and especially Steffi for their understanding.

Contents

1 General Introduction

1.1 Two-Sided Markets – Definition, Challenges and Examples

Over the last years, the Internet has substantially changed people's lives in many ways. Numerous interactions, which before the rise of the Internet were only possible offline, can now be carried out online as well. As transaction costs were reduced compared to traditional offline markets, new types of platforms were created. Not only are there online equivalents of traditional marketplaces, but additionally, new markets for information exchange and advertising such as online search engines and news portals have emerged.

These revolutionary changes have also sharpened the way economists think about markets and transactions in general. In particular, economists' attention was attracted to the specific characteristics of two-sided markets, which are defined as markets with platforms serving two different user groups that exert inter-group network effects on each other. Such inter-group network effects arise if on a given platform the utility for each user on one side changes ceteris paribus with the number of users on the other side.[1] Equivalently, Rochet and Tirole (2006, abstract) "identify two-sided markets with markets in which the structure, and not only the level of prices charged by platforms, matters." Together, these two alternative definitions reflect the essential feature that distinguishes two-sided markets from classical multi-product markets: the network effects between the two user groups imply a pricing system on the platform that can differ from that obtained for classical multi-product markets without inter-group network effects.

The inter-group network effects of two-sided markets have important managerial implications. Eisenmann et al. (2006) analyze potential business strategies for platform operators in two-sided markets. According to them, platform operators have to face three major challenges. First, they have to find the right pricing system for their platform. Second, they have to deal with the winner-take-all dynamics, which describes the fact that due to the inter-group network effects, two-sided markets often exhibit a tendency that the whole market is served by a single platform. Third, platform operators

[1] Network effects may also arise in classical one-sided markets. For a detailed discussion the reader is referred to section 1.2.

have to fight the threat of envelopments, i.e. the danger that other platforms may start to offer the service as part of a bundled product, thereby attracting users of one's own platform. In this context, Hagiu (2007) describes design and expansion strategies for multi-sided platforms.

The implications of the specific characteristics of two-sided markets in terms of antitrust economics are analyzed by Evans (2003). Wright (2004) studies the consequences of applying one-sided economic logic to markets that are actually two-sided and identifies several fallacies resulting from such a misleading analysis. In particular, he argues that since prices in two-sided markets need not equal marginal costs to be efficient, a high price-cost margin does not necessarily indicate market power and prices below marginal costs need not indicate predation, either. As a consequence, even in mature two-sided markets, price structures that do not reflect costs may be justified. Regulating prices set by a platform in a two-sided market is not competitively neutral and an increase in competition does not have to result in a more efficient or a more balanced price structure (see Wright, 2004).

For these reasons, the question to what extent a market is actually two-sided or not is crucial for both the platform operator and the policymaker. The more important the inter-group network effects in a two-sided market are, the bigger are the potential consequences with respect to the platform's profit but also to economic welfare if these network effects are ignored. Therefore, both platform operators and policymakers have an interest in that the inter-group network effects are identified and quantified as precisely as possible.

Two-sided markets already existed before the rise of the Internet. Classical examples in the offline world are manifold and include nightclubs, consumer magazines, yellow pages, payment card systems and telecommunications. However, the network character of the Internet implies a particularly strong tendency for the creation of platforms with different user groups and inter-group network effects. Specific examples of two-sided markets in the Internet comprise online search engines, news portals, online marketplaces, online dating platforms and online social lending platforms. More broadly, any website where some users provide and some other user consume content and/or advertising can be regarded as a two-sided market. Finally, even the Internet backbone itself is a two-sided market.

1.2 Related Literature

This dissertation is related to various strands of the theoretical and empirical literature on network externalities in general and on two-sided markets in particular. According to Katz and Shapiro (1985, p. 424), a good exhibits network externalities if “the utility that a user derives from consumption of the good increases with the number of other agents consuming the good”. The authors distinguish between direct and indirect network externalities. Direct network externalities occur “through a direct physical effect of the number of purchasers on the quality of the product” (Katz and Shapiro, 1985, p. 424) while indirect network externalities are related to complementary goods whose price and quantity vary with the number of users of the considered principal good. Katz and Shapiro (1985, p. 424) explain this with the example of hardware and software: “For example, an agent purchasing a personal computer will be concerned with the number of other agents purchasing similar hardware because the amount and variety of software that will be supplied for use with a given computer will be an increasing function of the number of hardware units that have been sold” (Katz and Shapiro, 1985, p. 424). Katz and Shapiro (1994), Economides (1996) and Shy (2001) provide excellent overviews on the literature on network effects.[2]

The definition of indirect, inter-group network effects leads directly to the formal analysis of two-sided markets, provided in two seminal papers by Armstrong (2006) and Rochet and Tirole (2006). These two papers differ from each other in the fact that Armstrong (2006) models users’ participation on the platform, whereas Rochet and Tirole (2006) analyze platform usage.[3] However, the main result of both papers is that when platform operators take into account the inter-group network externalities that characterize two-sided markets, then equilibrium prices differ from those in classical one-sided markets. Even under perfect competition prices on the two sides need not be equal to the respective marginal costs. Often, one side is subsidized and all profit is made on the other side, where demand may be more inelastic and/or where the network effects induced by the first side may be stronger.

Often, users in two-sided markets have the possibility to operate on several or even

[2]Liebowitz and Margolis (1994) distinguish between network effects and network externalities. They define network externalities as those network effects that imply a market failure. In what follows, I will use the two terms interchangeably; which term is appropriate according to the definition above depends on whether the model implies market failure or not (also see Grajek, 2004).

[3]More specifically, in the model by Rochet and Tirole (2006), the product of the number of users on the two sides represents the potential number of interactions on the platform. For each single interaction the total price is split between the two members who actually interact. In the model by Armstrong (2006), consumers on the two sides pay for the right to interact on the platform.

all platforms at the same time. In the literature, this feature is referred to as "multi-homing". The most interesting case occurs when users on just one side of the market are able to multi-home while users on the other side single-home. In this case, which is described as a "competitive bottleneck" in the literature, the single-homing side is treated favorably, while the interests of the multi-homing side are ignored (see Caillaud and Jullien, 2003; Armstrong, 2006; Rochet and Tirole, 2006). This contradicts the intuition that those who multi-home (and thus seem to be more flexible) should be better off. Dynamic analyses of two-sided market platforms are also particularly interesting because the inter-group network effects result in a "chicken-and-egg" problem: the platform operator needs to get users from both sides on board (see Caillaud and Jullien, 2003).

In addition to these general papers on two-sided markets, the other theoretical literature mostly investigates economic questions in the context of specific two-sided markets. One illustrative example is competition between different payment card systems, where the two user groups are card holders and merchants. The fee paid to the payment card system for network usage and/or for each single transaction is split up between these two groups and is known as the "interchange fee". In this context, Rochet and Tirole (2002), Schmalensee (2002), Gans and King (2003), Wright (2003), Rochet and Tirole (2008), and Hermalin and Katz (2006) provide formal analyses of payment systems. Another specific two-sided markets industry is telecommunications. Here, the two sides are the originators and the receivers of calls. Network competition and network interconnection are formally analyzed for example by Armstrong (1998), Wright (2002), and Jeon et al. (2004). Church and Gandal (2006) present several case studies on competition in telecommunications.

This dissertation is more closely related to two-sided markets in the media industries and in particular in the Internet. Anderson and Coate (2005) and Kind et al. (2005) provide theoretical analyses of television broadcasting to determine the socially optimal level of television advertising. Other theoretical papers study the print media, where the two market sides are readers and advertisers. Gabszewicz et al. (2001) and Ferrando et al. (2003) develop basic models for the analysis of this market. The specific features of two-sided markets in the Internet are studied by Caillaud and Jullien (2001) and Jullien (2005).[4] More precisely, Caillaud and Jullien (2001, p. 799) "account for some specificities of cybermediation, and in particular the pricing schemes and the possibility of multiple registration". This underlines the fact that multi-homing is of particular importance on Internet platforms. Katsamakas and Bakos (2004) study ownership and

[4]Related to this, Evans et al. (2005) and Economides and Katsamakas (2006) investigate two-sided markets in the context of operation systems and software.

investment decisions in two-sided Internet platforms and find “that, under standard assumptions about the design technologies, the network has a strong incentive to invest only on one of the sides and let the other side participate with a marginally positive investment” (Katsamakas and Bakos, 2004, p. 15). Ellison et al. (2004) analyze the scale effect in online auctions, for example on eBay, Yahoo! and Amazon.

While the theoretical literature on two-sided markets becomes increasingly developed, the empirical work remains relatively sparse as the measurement of network effects in two-sided markets is particularly challenging.[5] The inter-group network effects occur simultaneously and create a feedback loop in equilibrium. In order to account for this loop, the two sides of the market have to be considered together. It is possible that the inter-group network effects in one or both directions are negative (see Reisinger, 2004).[6] But even if network effects in both directions are positive, they can still be differently shaped and/or of different quantitative importance.

Separating these effects in the estimation procedure is not straightforward and requires a structural approach rather than a purely descriptive analysis.[7] For such a structural approach, extensive data is needed on prices, quantities, covariates and instrumental variables on the two sides and over the same observation period.[8]

Similar to the theoretical work on two-sided markets, a large share of the empirical literature concentrates on the media. Rysman (2004) builds a structural economic model to estimate network effects between consumers and advertisers in yellow pages directories. Similarly, Kaiser and Wright (2006) and Kaiser and Song (2009) quantify network effects in the magazines industry. One central finding of these papers is that due to the inter-group network effects, advertisers typically cross-subsidize readers. It is also shown that the actual amount of this subsidy need not be socially optimal. Argentesi and Filistrucchi (2007) estimate market power in the newspaper industry by comparing the observed price structure with that of different models of coordinated price setting.

[5] There are also several papers on the measurement of direct network effects in classical one-sided markets (see e.g. Economides and Himmelberg, 1995). Yet, this is not the focus of this dissertation.

[6] A typical example is television advertising. In the models by Anderson and Coate (2005) and by Kind et al. (2005), the utility of an advertiser increases ceteris paribus as the number of viewers increases (positive network effects from viewers to advertisers), whereas viewers dislike television advertisements (negative network effects from advertisers to viewers).

[7] As Reiss and Wolak (2007) point out in their overview of structural econometric modeling in industrial organization, the difference between descriptive and structural analyses is that the latter approach builds on an empirical model derived from underlying econmic and statistical assumptions. Therefore, the estimation of a structural model makes it possible to derive conclusions about specific economic primitives.

[8] As pointed out in the second paper of this dissertation, under certain conditions a semi-structural analysis also enables the econometrician to estimate the network effects in a two-sided market.

1.3 Contribution of this Dissertation

As outlined above, the specific challenges that arise in two-sided markets for both platform operators and policymakers are diverse and make the economic analysis particularly interesting. Among others, these challenges comprise optimal platform pricing, platform dynamics and – in a broad sense – the question how potential platform intermediaries have to be rewarded such that adverse incentives for platform members are avoided. In this context, this dissertation analyzes various aspects of two-sided markets and contributes to the existing literature described above in the following different ways.

In the first paper, I study the interactions between the traditional (offline) demand channels and the new (online) demand channels in the German magazines industry. In particular, I focus on the substitution between offline and online advertising, which has not yet been studied either in the theoretical literature or empirically. Yet, this interaction becomes increasingly important as with the rise of the Internet, traditional consumer magazines create companion websites and make much of the contents available online to readers, which in turn attracts advertisers to these websites. The degree of substitution between offline and online advertising describes how easy or difficult it is for advertisers to reallocate their advertising budget from the traditional consumer magazines to the new companion websites. This question is essential for platform operators seeking to maximize joint profit from the traditional magazine and the companion website. Additionally, the paper also contributes to the general theoretical literature on two-sided markets. While the model developed in this paper is adapted to the case of consumer magazines and their companion websites, it can be modified to fit other two-sided markets where interactions between different demand channels shall be investigated.

In the second paper of this dissertation, I propose a method to estimate network effects in potentially two-sided markets absent data on prices and quantities. Somewhat similar semi-structural reduced-form revenue regressions are used by Panzar and Rosse (1987) to derive conclusions on market power in classical one-sided markets, yet this paper is the first one to apply this methodology to two-sided markets. Compared to existing fully structural techniques this approach has the advantage that the resulting regression equations can be easily estimated via OLS. Moreover, neither data on prices or quantities nor instrumental variables are needed: only revenue data and demand and/or cost shifters are used.

The third paper of this dissertation sheds light on a different aspect of two-sided markets. In particular, it investigates the certification mechanisms and incentives that enable lending markets to match demand and supply despite the absence of traditional financial intermediaries with skin in the game. This issue has gained particular attention with the financial and economic crisis, which has partly been the consequence of irresponsible financial intermediation and disintermediation. However, its analysis is typically difficult because of the lack of adequate data. In this context, we analyze the online social lending market, in which potential lenders and borrowers interact directly without a financial intermediary but can create self-organized groups instead. We show how improperly designed incentives for the group leaders may lead to adverse outcomes for the market as a whole. Despite the specific characteristics of this online social lending market the results also translate to other lending markets and comprise important general lessons for consumer lending.

2 Multi-Channel Demand in Two-Sided Markets: Evidence from the Magazines Industry

2.1 Introduction

Traditionally, consumer magazines have operated in a two-sided market, serving two distinct groups of customers at the same time: readers and advertisers. The market is a two-sided one because these two groups exert inter-group network effects on each other: in a given magazine, each advertiser ceteris paribus values each additional reader, while each reader (dis-)likes each additional advertisement.

However, with the rise of the Internet, further demand channels are explored: readers can now access much of the contents also on the magazines' websites, which in turn attracts advertisers to these websites. In recent years, online advertising has gained more and more importance, as advertisers have shifted their budgets from the traditional media – in particular from consumer magazines – to the Internet.

In order to better understand this shift, in this paper I investigate the interactions between the traditional (offline) demand channels and the new (online) demand channels in the German magazines industry. More precisely, the contribution of this paper is to answer the question how demand for magazines and their companion websites is interrelated on both the side of the readers and the side of the advertisers. In particular, given the enormous shift from offline to online advertising described above, I expect to find that these two types of advertising are substitutes. In contrast, had offline advertising and online advertising experienced a simultaneously increasing trend with the rise of the Internet over the last years, it would be more intuitive to expect a complementarity between the two advertising types.

The interaction on the side of the advertisers has not yet been studied either in the theoretical literature or empirically but has potentially important managerial implications for both advertisers and platform operators. First, the higher the degree of substitution between offline and online advertising, the easier it is for advertisers to reallocate their advertising budgets from traditional offline advertising to online advertising. Since advertisers and readers operate in a two-sided market, this reallocation will then also lead to a change in the number of readers on the respective other side of the market. Second, in order to maximize total platform profit from advertisers and readers both in the offline and in the online world, each platform operator needs to take these additional interactions into account when setting prices.

The data I use is particularly rich in many dimensions. Apart from prices and quantities on offline and online advertising and reading, I also possess information on magazine composition, website composition as well as magazine readers' and website viewers' characteristics. While this rich dataset – together with adequate modeling and appropriate instrumental variables – makes structural estimation in principle possible, there is one important shortcoming: The precise measurement of readers' and advertisers' actions on the websites has started only recently. Therefore, not all data is available over the entire time period of interest. This calls for a special estimation procedure: I first construct a demand model in which there is only the offline world and estimate it, using the General Method of Moments (GMM), with data from the time period when neither online reading nor online advertising played a significant role. Under the assumption that with the rise of the Internet the estimated coefficients of this nested model do not change, I then estimate the remaining parameters of the full model, in which offline and online advertising and reading coexist.

This study is related to three different strands of the literature. First, it contributes to the theoretical literature on (indirect) network externalities in general and on two-sided markets in particular. Two-sided markets are formally analyzed in the two seminal papers by Armstrong (2006) and Rochet and Tirole (2006). Roson (2005) provides an early survey on the economics of two-sided markets.[1] In this general context, this paper is the first one that explicitly models demand interactions between different two-sided market demand channels: On the side of the readers, these two channels are offline and online reading; on the side of the advertisers these are offline and online advertising. Modeling these interactions across different demand channels becomes increasingly important in the economic analysis not only in this particular setting of consumer magazines, but also in two-sided markets in general. With the rise of the Internet, many online two-sided platforms have emerged, and as a consequence both consumers and firms face an increasing number of possibilities to interact with each other. Ignoring substitution or complementarity effects between the corresponding demand channels of different plat-

[1] For a more detailed discussion on related work on two-sided markets, the reader is referred to the literature review in section 1.2 of this dissertation.

forms may lead to potentially biased economic conclusions by platform users.

The second strand of literature comprises the growing number of empirical papers on the estimation of inter-group network effects in two-sided markets, in particular in media industries. Wright (2004) argues that ignoring network externalities may lead to wrong conclusions with respect to the policy implications to undertake for regulation in a two-sided market. Therefore, quantifying these network externalities is of great importance. Various researchers have estimated network effects in television (e.g. Anderson and Coate, 2005; Kind et al., 2005), newspapers (e.g. Argentesi and Filistrucchi, 2007), magazines (e.g. Kaiser and Wright, 2006; Kaiser and Song, 2009) or yellow pages (e.g. Rysman, 2004). Generally, all these papers find that there are positive network effects from consumers to advertisers: advertisers like consumers. The estimation results of network effects in the opposite direction – from advertisers to consumers – vary: most papers assume a priori that consumers dislike advertisements (e.g. Anderson and Coate, 2005), whereas others find that consumers like advertisements (e.g. Rysman, 2004; Kaiser and Song, 2009). The specific research design of my study allows me to estimate network effects in the two-sided market of the German magazines industry in the absence of the online world. In doing so, I am able to reproduce the empirical results found in the literature in my first estimation step, although with a slightly different model.

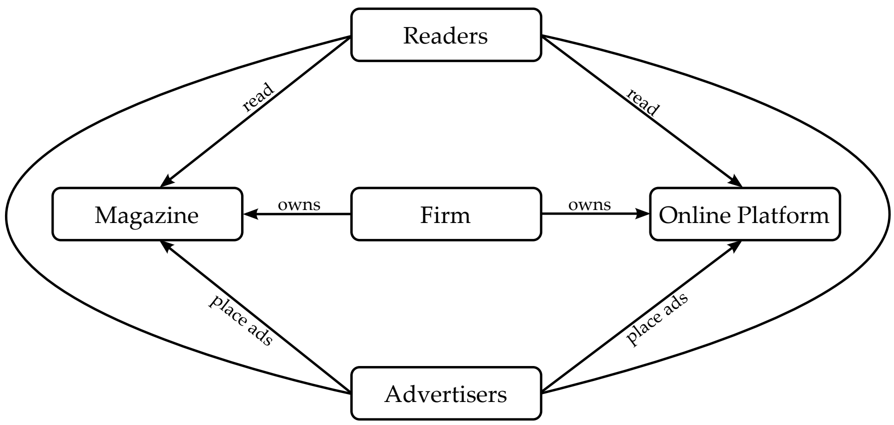
Finally, this paper contributes to the ongoing debate of whether the rise of the Internet has influenced demand for reading and demand for advertising in the traditional magazines and newspaper industry. Kaiser (2002) finds that in the German women's magazines market observed between 1998 and 2001, website provision neither has a significant effect on magazine nor on advertising demand, whereas Kaiser (2006) provides empirical evidence in favor of the belief that a magazine's companion website induces channel competition on its print version. In his study of the Italian newspaper market, Filistrucchi (2005) finds that opening a website has a negative impact both on the sales of the newspaper who opens it and on those of its rivals. Gentzkow (2007) develops a model that allows for complementarity between offline and online newspapers to study competition between print and online newspapers, and estimates the relationship between the print and online papers in demand, the welfare impact of the online paper's introduction, and the expected impact of charging positive online prices. In this context, my paper is the first one that does not only consider the interactions on the side of the readers, but also those on the side of the advertisers by explicitly modeling and estimating the interaction between offline and online advertising. Ignoring this interaction may lead to a suboptimal choice of the amount of advertising both offline and online. Therefore, this latter contribution has important managerial implications for the advertisers seeking to optimize joint profit from their offline and online advertising channels.

This paper is structured as follows: In the next section, I construct a model of two-sided markets with different demand channels that is suitable for estimation in two steps as described above. Section 2.3 describes the industry and the data. In section 2.4, I present and discuss the estimation results. Section 2.5 concludes.

2.2 The Model

In this section, I construct a model of (inverse) demand for advertising and demand for reading that captures the two-sidedness of both markets as well as the interrelation between the offline and the online channel for both advertising and reading.

Figure 2.1: Model Setup – Full Model



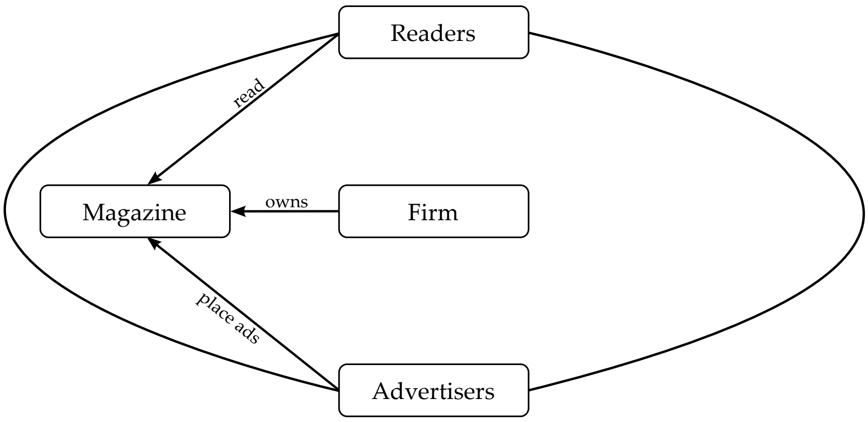
As shown in figure 2.1, I consider a number of firms $j = 1, \ldots, J$, each of which owns a magazine (m) and a website (w). Each firm maximizes joint profit from its magazine and its website. Readers can read the magazine offline, but also surf over the corresponding website in order to obtain the same or complementary information.[2] Advertisers can place advertisements in the magazine, but also – in the form of banners – on the corresponding website. Therefore, for the individual firm j, there are three sources of revenue: revenue from readers of the magazine, revenue from advertisements in the magazine and revenue from advertisements on the compantion website. There is no price charged to consumers for surfing over the website.

[2]The question whether the same or complementary information can be obtained from the magazine and its compantion website has direct implications for the question whether these two products are substitutes or complements. Unfortunately, my data does not provide this information.

Readers and advertisers operate in a two-sided market, i.e. they exert inter-group network effects on each other: for example, offline advertisers value each additional reader of the magazine, and magazine readers' utility probably changes with the number of advertisements in the magazines. These network effects are represented by the outer arrows from readers to advertisers and vice versa in figure 2.1.

As indicated in the introduction, estimation of this model proceeds in two steps, because not all data is available for the entire time period of interest:[3] I first consistently estimate a model in which there is only the offline world using data from a time period when Internet reading and online advertising did not play a significant role. This model is shown in figure 2.2. In the second step and under the assumption that the estimated coefficients of the offline world do not change, I then insert the estimates I obtain from the first step into the full model where there is both the offline and the online world (shown in figure 2.1) and estimate the remaining coefficients.[4] To make the two-step estimation possible, the model in which only the offline world exists (see figure 2.2) has to be nested in the full model (see figure 2.1).

Figure 2.2: Model Setup – Nested Model



[3]Parker and Röller (1997) use a somewhat similar method to determine to what degree duopolistic competition leads to competitive market outcomes. In particular, they also consider a two-period model: a monopoly period followed by a duopoly one. This makes a specification test of their model for the monopoly period possible. Moreover – and this is also the case for my approach as discussed later – this increases the efficiency of the estimator for the second period.

[4]A detailed discussion of the advantages and disadvantages of this approach is delayed until section 2.4.3.

2.2.1 (Inverse) Demand for Advertising

A Model of (Inverse) Demand for Advertising

Consider a representative, price-taking advertiser who, in his decision how to choose offline and online advertisements on the different platforms, faces the following profit function:[5]

$$\Pi = \sum_{j=1}^{J} \Bigg[\underbrace{R(\bar{\pi}_{jm}, \bar{\pi}_{jw}, a_{jm}, a_{jw}, N_{jm}, N_{jw})}_{\text{Revenue from advertising on platform } j} - \underbrace{(p_{jm}a_{jm} + p_{jw}a_{jw})}_{\text{Costs from advertising on platform } j} \Bigg] \tag{2.1}$$

Here, N_{jm} and N_{jw} capture the number of readers and the number of page impressions on platform j, respectively. a_{jm} is the number of advertisement pages the representative advertiser chooses to place in the magazine of platform j and a_{jw} is the number of advertisement banners and their size (together measured by ad impressions) the representative advertiser chooses to place on the website of platform j. p_{jm} and p_{jw} are the corresponding prices for an offline and an online ad on platform j.[6] $\bar{\pi}_{jm}$ and $\bar{\pi}_{jw}$ scale the revenue per reader offline and website viewer online (i.e. how much revenue the representative advertiser can make per ad viewer offline ($\bar{\pi}_{jm}$) and online ($\bar{\pi}_{jw}$).

I assume that $R(\cdot)$ takes the form of a Constant Elasticity of Substitution (CES) function $R(\bar{\pi}_{jm}, \bar{\pi}_{jw}, a_{jm}, a_{jw}, N_{jm}, N_{jw}) = \left[\bar{\pi}_{jm}\left(a_{jm}^{\gamma_m} N_{jm}^{\beta_m}\right)^{\rho} + \bar{\pi}_{jw}\left(a_{jw}^{\gamma_w} N_{jw}^{\beta_w}\right)^{\rho}\right]^{\frac{1}{\rho}}$ and that there is separability between advertising on different platforms (see Rysman, 2004), but not between offline and online advertising on a given platform. This assumption is a rather strong one. Rysman (2004) shows that it follows directly from two other underlying assumptions. The first one is that readers do not use various directories at the same time. The second one is that advertisers' profit per consumer look is constant. Under these assumptions, the optimization problem of the representative advertiser, which is derived from equation (2.1), simplifies to the choice of a_{jm} and a_{jw} on any given platform j:

$$\max_{a_{jm}, a_{jw}} \left[\bar{\pi}_{jm}\left(a_{jm}^{\gamma_m} N_{jm}^{\beta_m}\right)^{\rho} + \bar{\pi}_{jw}\left(a_{jw}^{\gamma_w} N_{jw}^{\beta_w}\right)^{\rho}\right]^{\frac{1}{\rho}} - (p_{jm}a_{jm} + p_{jw}a_{jw}) \tag{2.2}$$

[5] In this subsection, I extend the model for advertising in yellow papers by Rysman (2004) in my theoretical specification of advertisement demand.

[6] In addition to the quantities a_{jm} and a_{jw}, these prices p_{jm} and p_{jw} may also be endogenous in the optimization problem of the representative advertiser – depending on the type of competition in the market: If there is perfect competition in both advertising markets, prices are completely exogenous to the representative advertiser. Otherwise, prices are endogenous and need to be instrumented for in the estimation since I do not model the supply side.

There are several economic assumptions underlying the functional specification made in equation (2.2). First, offline and online revenue are considered as substitutes or complements by the representative advertiser and $\sigma = 1/(1-\rho)$ measures the corresponding elasticity of substitution. More specifically, if $\rho = 1$, then offline and online advertising revenue are considered as perfect substitutes by the representative advertiser, for $\rho \to 0$, I get the Cobb-Douglas function and for $\rho \to -\infty$, I get the Leontief (perfect complements) function. The main aim of this study is to estimate the coefficient ρ.

γ_m and γ_w measure the curvature of revenue of the representative advertiser with respect to offline and online advertising. Put differently, they are closely linked to the own-price elasticities of offline and online advertising. I expect that $\gamma_m < 1$ and $\gamma_w < 1$, i.e. downward sloping demand curves for both offline and online advertising.

The functional specification (2.2) does not explicitly model a congestion effect of advertisements, i.e. the ability to reach consumers in a given medium is not explicitly assumed to vary with the number of competing advertisers. If this is the case after all, the profit the representative advertiser can make on a given platform j will also depend on total advertisement offline or online on this platform. Rysman (2004) accounts for this congestion effect in his model but is unable to disentangle it from the decreasing returns to the individual advertiser from large advertisements as only the sum of the two effects is identified. Therefore, if a congestion effect is indeed present although it is not specified in the economic model (2.2), then γ_m (respectively γ_w) will capture this congestion effect in addition to the curvature effect. In this case, the estimate of the curvature effect will be biased downward.

Network effects from readers and website viewers to the representative advertiser are captured by β_m and β_w. In line with previous research, I expect that $0 < \beta_m < 1$, indicating positive non-explosive network effects from readers to advertisers, i.e. – all else equal – each additional reader increases the representative advertiser's profit. Similarly, $0 < \beta_w < 1$ would imply positive network effects from website surfers to advertisers in the online market.

From the maximization problem (2.2), the first order conditions can be derived as follows:

$$\begin{aligned}
\frac{\partial \Pi}{\partial a_{jm}} &= 0 \Leftrightarrow p_{jm} = \bar{\pi}_{jm}\gamma_m a_{jm}^{\gamma_m\rho-1} N_{jm}^{\beta_m\rho}\left[\bar{\pi}_{jm}\left(a_{jm}^{\gamma_m}N_{jm}^{\beta_m}\right)^{\rho} + \bar{\pi}_{jw}\left(a_{jw}^{\gamma_w}N_{jw}^{\beta_w}\right)^{\rho}\right]^{\frac{1-\rho}{\rho}} \\
\frac{\partial \Pi}{\partial a_{jw}} &= 0 \Leftrightarrow p_{jw} = \bar{\pi}_{jw}\,\gamma_w a_{jw}^{\gamma_w\rho-1}\, N_{jw}^{\beta_w\rho}\left[\bar{\pi}_{jm}\left(a_{jm}^{\gamma_m}N_{jm}^{\beta_m}\right)^{\rho} + \bar{\pi}_{jw}\left(a_{jw}^{\gamma_w}N_{jw}^{\beta_w}\right)^{\rho}\right]^{\frac{1-\rho}{\rho}}
\end{aligned} \tag{2.3}$$

Assuming that there are K equal advertisers in the market, in a symmetric equilibrium one can replace $a_{jm} = A_{jm}/K$ and $a_{jw} = A_{jw}/K$ and obtains after simplifying the first order conditions (2.3):[7]

$$
\begin{aligned}
p_{jm} &= \bar{\pi}_{jm}^{\frac{1}{\rho}} K^{1-\gamma_m} \gamma_m A_{jm}^{\gamma_m - 1} N_{jm}^{\beta_m} \underbrace{\left[1 + \frac{\bar{\pi}_{jw} K^{(1-\gamma_w)\rho}}{\bar{\pi}_{jm} K^{(1-\gamma_m)\rho}} \left(\frac{A_{jw}^{\gamma_w}}{A_{jm}^{\gamma_m}} \frac{N_{jw}^{\beta_w}}{N_{jm}^{\beta_m}} \right)^{\rho} \right]^{\frac{1-\rho}{\rho}}}_{\substack{\text{Term due to substitution or complemen-} \\ \text{tarity between offline and online world}}} \\
p_{jw} &= \bar{\pi}_{jw}^{\frac{1}{\rho}} K^{1-\gamma_w} \gamma_w A_{jw}^{\gamma_w - 1} N_{jw}^{\beta_w} \underbrace{\left[1 + \frac{\bar{\pi}_{jm} K^{(1-\gamma_m)\rho}}{\bar{\pi}_{jw} K^{(1-\gamma_w)\rho}} \left(\frac{A_{jm}^{\gamma_m}}{A_{jw}^{\gamma_w}} \frac{N_{jm}^{\beta_m}}{N_{jw}^{\beta_w}} \right)^{\rho} \right]^{\frac{1-\rho}{\rho}}}_{\substack{\text{Term due to substitution or complemen-} \\ \text{tarity between offline and online world}}}
\end{aligned}
\tag{2.4}
$$

Let $\pi_{jm} = \bar{\pi}_{jm} K^{(1-\gamma_m)\rho}$ (respectively $\pi_{jw} = \bar{\pi}_{jw} K^{(1-\gamma_w)\rho}$) account for both the number of advertisers in the market and the advertising revenue per reader (respectively website viewer), such that (2.4) simplifies to:

$$
\begin{aligned}
p_{jm} &= \pi_{jm}^{1/\rho} \gamma_m A_{jm}^{\gamma_m - 1} N_{jm}^{\beta_m} \underbrace{\left[1 + \frac{\pi_{jw}}{\pi_{jm}} \left(\frac{A_{jw}^{\gamma_w}}{A_{jm}^{\gamma_m}} \frac{N_{jw}^{\beta_w}}{N_{jm}^{\beta_m}} \right)^{\rho} \right]^{\frac{1-\rho}{\rho}}}_{\substack{\text{Term due to substitution or complemen-} \\ \text{tarity between offline and online world}}} \\
p_{jw} &= \pi_{jw}^{1/\rho} \gamma_w A_{jw}^{\gamma_w - 1} N_{jw}^{\beta_w} \underbrace{\left[1 + \frac{\pi_{jm}}{\pi_{jw}} \left(\frac{A_{jm}^{\gamma_m}}{A_{jw}^{\gamma_w}} \frac{N_{jm}^{\beta_m}}{N_{jw}^{\beta_w}} \right)^{\rho} \right]^{\frac{1-\rho}{\rho}}}_{\substack{\text{Term due to substitution or complemen-} \\ \text{tarity between offline and online world}}}
\end{aligned}
\tag{2.5}
$$

Note that for $A_{jw} = 0$ and $\rho = 1$, which are reasonable assumptions for the only-offline world, the additional term due to substitution or complementarity in the first equation in (2.5) disappears and the equation collapses into an estimation equation of a model where only the offline-world exists:

$$
p_{jm} = \pi_{jm} \gamma_m A_{jm}^{\gamma_m - 1} N_{jm}^{\beta_m}. \tag{2.6}
$$

Equation (2.6) is precisely the equation of inverse demand for advertising from the

[7] I assume that all advertisers are active in both the market for offline advertising and the market for online advertising.

model developed by Rysman (2004), with the exception that – as described above – I do not consider a congestion effect in advertising in my model.

Identification

For estimation of equation (2.6), i.e. of (inverse) demand for advertising in the nested model, I assume that $\ln \pi_{jm} = X_{jm}\delta_m + \nu_{jm}$, where $X_{jm}\delta_m$ is a linear combination of observable variables and ν_{jm} is an unobservable term. Taking the natural logarithm of equation (2.6) leads to:

$$\ln p_{jm} = (\gamma_m - 1) \ln A_{jm} + \beta_m \ln N_{jm} + \ln \gamma_m + X_{jm}\delta_m + \nu_{jm}. \tag{2.7}$$

As Rysman (2004) argues, the equilibrium quantity of advertising depends on price, and therefore A_{jm} and N_{jm} are potentially correlated with ν_{jm} and need to be instrumented for. In GMM, instrumental variables are applied to entire equations, not to single potentially endogenous variables. Essentially, good instruments for the estimation of (2.7) should be uncorrelated with the error term ν_{jm}, but correlated with advertising A_{jm} and/or reading N_{jm}.

As Rysman (2004) points out, ideal instruments for A_{jm} are variables that shift marginal cost and therefore the publishers' first order conditions. Accordingly, I use producer price index of printing material as well as wages paid in printing and publishing.[8] I also use the logarithm of GDP and firms' perceptions on their current business situation, which is captured by the Ifo Business Climate Index. These two variables reflect the overall situation of the economy, and I expect them to shift the publishers' first order conditions without shifting advertising demand at the same time.

As instruments for N_{jm}, i.e. the number of people who read a magazine, I use the unemployment rate, assuming that when more people are unemployed, either less magazine copies are sold because people have less money to spend or more copies are sold because people have more time to read.[9] Furthermore, I use the consumer price index as an instrument for N_{jm}. For this instrument, the assumption is that this global measure of changes in the price level of consumer goods and services purchased by households also shifts reading.

For estimation of the remaining coefficients in the advertising equations of the full

[8] Additionally, I experiment with producer price indices of printing machines and paper machines. However, these instruments work less well, probably since investments into printing machines are not made on a regular basis.

[9] Which of these two opposing effects prevails, is not important for the instrument to work well.

model (2.5), I specify:

$$
\begin{aligned}
\ln p_{jm} = {} & (\gamma_m - 1)\ln A_{jm} + \beta_m \ln N_{jm} + \frac{1-\rho}{\rho}\ln\left[1 + \frac{X_{jw}\delta_w}{X_{jm}\delta_m}\left(\frac{A_{jw}^{\gamma_w}}{A_{jm}^{\gamma_m}}\frac{N_{jw}^{\beta_w}}{N_{jm}^{\beta_m}}\right)^{\rho}\right] \\
& + \ln\gamma_m + \frac{1}{\rho}X_{jm}\delta_m + \nu_{jm} \\
\ln p_{jw} = {} & (\gamma_w - 1)\ln A_{jw} + \beta_w \ln N_{jw} \;\; + \frac{1-\rho}{\rho}\ln\left[1 + \frac{X_{jm}\delta_m}{X_{jw}\delta_w}\left(\frac{A_{jm}^{\gamma_m}}{A_{jw}^{\gamma_w}}\frac{N_{jm}^{\beta_m}}{N_{jw}^{\beta_w}}\right)^{\rho}\right] \\
& + \ln\gamma_w + \frac{1}{\rho}X_{jw}\delta_w + \nu_{jw}.
\end{aligned}
\tag{2.8}
$$

Recall that by assumption, γ_m, β_m as well as δ_m are consistently estimated in the first estimation step already. These estimates are assumed not to change over time and are therefore plugged into the system of equations (2.8). Consequently, in the second estimation step, I only need to instrument for the remaining variables that are potentially correlated with ν_{jm} in the first equation of (2.8) and/or with ν_{jw} in the second equation of (2.8): $\ln A_{jw}$ and $\ln N_{jw}$. As instruments for these two variables, I use the income distribution of the website viewers.

2.2.2 Demand for Reading

A Model of Demand for Reading

I assume that the demand for reading offline and online on platform j follows the log-linear specification:

$$
\begin{aligned}
\ln N_{jm} &= \phi_m \ln A_{jm} + \phi_w \ln A_{jw} + \psi_w \ln N_{jw} + \kappa_m cp_{jm} + Y_{jm}\theta_m + \epsilon_{jm} \\
\ln N_{jw} &= \chi_w \ln A_{jw} + \chi_m \ln A_{jm} + \tau_m \ln N_{jm} + Y_{jw}\theta_w + \epsilon_{jw},
\end{aligned}
\tag{2.9}
$$

where ϕ_m measures the network effects from offline advertisements on offline reading and ϕ_w describes the influence of online advertisements on offline reading (expected to be insignificant). cp_{jm} is the copy price of the magazine of platform j whereas there is no price for reading online in my model.[10] ψ_w and τ_m measure whether magazines and websites are substitutes (ψ_w,$\tau_m < 0$) or complements (ψ_w,$\tau_m > 0$). χ_m captures network effects from online advertisers to website viewers and χ_w measures the influence

[10] Depending on competition in the market, cp_{jm} may be endogenous and in this case needs to be instrumented for.

of offline advertisements on online reading. Y_{jm} (respectively Y_{jw}) captures magazine-specific (respectively website-specific) covariates of platform j.

I employ the log-linear specification for offline reading in the first equation of (2.9) instead of a (nested) logit one for two reasons:[11] First, since I analyze magazines from very different genres, the assumption that readers single-home is not justified. Second, for a (nested) logit specification I would need to observe the whole market in all periods in order to be able to compute the outside option consistently. This, however, is not the case for my data. Since users do not surf over just one, but over many websites, a nested logit model is not appropriate for the specification of demand for online reading either.

Similar to the case of advertising demand, for the reasonable assumptions for the only-offline world $\ln A_{jw} = 0$ and $\ln N_{jw} = 0$, the first equation of (2.9), which quantifies demand for offline reading, collapses to:

$$\ln N_{jm} = \phi_m \ln A_{jm} + \kappa_m cp_{jm} + Y_{jm}\theta_m + \epsilon_{jm}. \tag{2.10}$$

Identification

In equation (2.10), advertising A_{jm} and copy price cp_{jm} are potentially endogenous as they are simultaneously determined with reading N_{jm}. As instruments for both of these variables, I use the instrumental variables described in section 2.2.1 that are expected to drive marginal costs and therefore to shift the first order conditions of the publishers: producer price indices of printing material, wages paid in printing and publishing, GDP and firms' perceptions on their current business situation (Ifo Business Climate Index). Furthermore, I also use the income distribution of the magazines' readers and the average number of advertisement pages in all other magazines as instruments.

Similar to advertising demand, in the second estimation step only those coefficients have to be estimated in (2.9) that have not been estimated already in the first one. This means, that ϕ_m, κ_m and θ_m have not to be estimated in the second step. As a consequence, in the first equation of (2.9), I do not need to instrument for $\ln A_{jm}$ and cp_{jm}. In the second equation of (2.9), I need to instrument for $\ln A_{jw}$. As instrumental variables I use the income distribution of the website viewers, expecting that their income has an impact on advertising demand without influencing the number of page impressions the website generates.

[11] Kaiser and Song (2009) employ a similar log-linear model for reading demand.

2.3 Industry Characteristics and Data

2.3.1 Industry Characteristics

The magazines industry has a long tradition in Germany, and for a long time magazines have been the third largest advertising medium in Germany after television and newspapers (see Circle of Online Marketers, 2010). In sharp contrast to that, online reading and online advertising are relatively new markets, which have grown fast with the rise of the Internet. This contrast makes an analysis of the interactions in these two closely related and yet so different industries particularly interesting.

Figure 2.3 shows how the situation of the average magazine evolves over time. It presents the quarterly average number of readers and the quarterly average number of advertisement pages per magazine since 1990. Overall, the number of readers decreases, in particular from the year 1998 on. One possible explanation for this additional decrease might be that at that time more and more magazines created a companion website.[12] The average number of advertisement pages is remarkably stable until the year 2001 and then decreases considerably. While one has to be careful when inferring causality from this graph, this implies at least a strong positive correlation between the average number of copies sold and the average number of advertisement pages and suggests the existence of inter-group network effects between reading and advertising.

A somewhat similar picture can be inferred from figure 2.4. Average revenue from readers decreases from 1990 on, in particular since 2000. In contrast, average revenue from advertisers slightly increases until 2000 and strongly decreases afterwards.

During the same time period, online advertising gains signifcant importance. In 2009, online advertising increases in Germany by 12.4% to reach a gross volume of more than four billion Euros, and surpasses with 16.1% advertising from consumer magazines (13.1%) for the first time in history (see Circle of Online Marketers, 2010).[13]

Together, these facts suggest that with the rise of the Internet both magazine readers and offline advertisers have to face changes in their interconnected markets. It is therefore essential to take into account the growing markets for both online reading and online advertising when analyzing these classical markets, in particular from the year 2000 on.

[12] Although not shown in the graph, the average number of page impressions of the corresponding websites increases considerably between 2002 and 2009.

[13] In this paper I only consider online advertising on the companion websites of the magazines, not on all websites.

Figure 2.3: Average Number of Readers and Average Number of Advertisement Pages

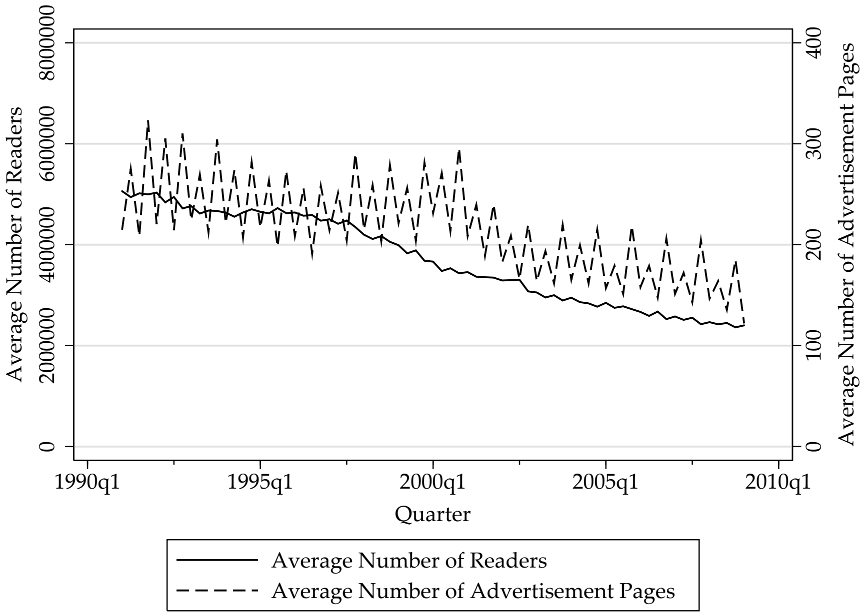
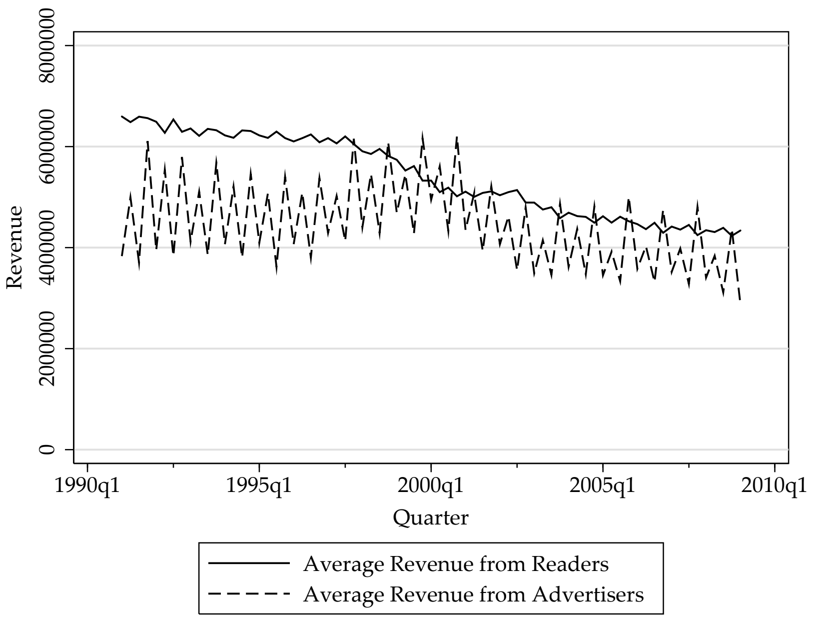


Figure 2.4: Average Revenue from Readers and Average Revenue from Advertisers



2.3.2 Data

The data is compiled from various sources.[14] The number of sold magazine copies, copy prices, number of advertising pages per magazine and price per advertising page is publicly available for download from *Focus Medialine* (http://medialine.focus.de) and was originally collected by the *Informationsgemeinschaft zur Feststellung der Verbreitung von Werbeträgern (IVW)*, a non-for profit institution that has been monitoring the dissemination of advertising media in Germany since 1949.[15] This data is available from 1972 until 2009 on a quarterly basis.

Newspaper composition in terms of content pages is analyzed regularly by the *Jahreszeiten Verlag* (a big German publisher) in its *Funktionsanalyse*. For each magazine and each year, this data set contains the number of pages in different content types. I aggregate this very detailed information, which is available from 1994 to 2009 with the exception of 1995, to six different types of pages (Fashion, Do-it-yourself, Family & Health, Travel & Hobby, Knowledge & Entertainment, Program & Service) and use the average over the available years as magazine-fixed effects.

The number of page impressions were provided by *IVW* and the number of ad impressions by *Nielsen Media Research*. Ad impressions are available on a monthly basis, yet only for 2009, which makes the two-step estimation employed in this paper necessary. A page impression is defined as a call of a website that is unique and induced by the user, an ad impression is defined accordingly.[16]

Data on magazine reader characteristics come from the *Arbeitsgemeinschaft Media-Analyse (AG.MA)*. In an extensive survey, *AG.MA* regularly interviews more than 25,000 people over the age of 14 on their reading habits. From this survey, I aggregate readers' characteristics for each magazine in the different years.

Website viewer characteristics – which are also compiled from a survey – and online advertisement prices were provided by the *Arbeitsgemeinschaft Online Forschung*

[14]Parts of this data is also used and described by Kaiser (2002).

[15]In line with Kaiser (2002), the prices I use in my analysis are computed as weighted (by the respective numbers of advertising pages) averages of black and white, two-color and four-color advertisement prices.

[16]According to IVW (2009), the action is "induced by the user" if he expects to obtain a significant change in the page contents. IVW gives some examples for such actions: This may be the call of a new website or new parts of a website or reloads of the same website or its parts as a response to a mouse click or to the use of the keyboard by the user. In contrast to that, actions which are not induced by the user are for example a call of a new website or new parts of a website by automatic forwarding, an automatic reload of the same website or its parts (e.g. newsticker), the call of a website when closing a window and the call of a website by robots or spiders. A "significant" change is for example a change in text passages or visual / multimedial contents that are contentwise in the center of the website, a new question in a quiz or a new picture in a slideshow.

(AGOF). This data is available on a quarterly basis. For my analysis, I use data from the year 2009.

The instrumental variables other than those from the sources mentioned above (i.e. GDP, Consumer Price Index, etc.) come from the German Federal Statistical Office and the Ifo Institute for Economic Research and are available on a quarterly basis.

For the first estimation step, I obtain a final sample of 3,963 observations of 137 magazines for the time period from the first quarter of 1991 until the fourth quarter of 1999. Table 2.1 provides univariate statistics on the endogenous and exogenous variables of reading and advertising as well as on the instrumental variables during this time first sample period.

Table 2.1: Univariate Statistics: Nested Model (Only Offline World)

	Mean	S.D.	Min	Max	N
PANEL A: ENDOGENOUS VARIABLES: OFFLINE ADS					
$\ln p_{jm}$	9.39	0.65	7.60	10.92	3,963
$\ln A_{jm}$	5.15	0.89	2.48	7.79	3,963
PANEL B: ENDOGENOUS VARIABLES: OFFLINE READING					
cp_{jm}	2.27	1.34	0.41	7.57	3,963
$\ln N_{jm}$	14.50	1.41	11.23	17.52	3,963
PANEL C: EXOGENOUS VARIABLES: OFFLINE ADS AND OFFLINE READING					
Fashion (%)	8.73	15.87	0.00	87.61	3,963
Do-it-yourself (%)	8.39	20.90	0.00	84.47	3,963
Family & Health (%)	9.17	12.38	0.00	65.37	3,963
Travel & Hobby (%)	20.31	24.95	0.10	89.04	3,963
Knowledge & Entertainment (%)	39.25	28.25	1.47	90.64	3,963
Program & Service (%)	14.13	19.44	3.30	86.21	3,963
ln(# Pages)	6.52	0.65	4.56	8.46	3,963
PANEL D: INSTRUMENTAL VARIABLES					
Wage: Printing	11.93	0.95	10.16	13.00	3,963
Producer Price Index: Paper	91.13	2.38	86.53	97.07	3,963
$\ln GDP$	4.50	0.04	4.44	4.58	3,963
Consumer Price Index	86.22	4.81	74.40	91.70	3,963
Ifo-Index	96.27	4.53	86.43	107.10	3,963
Unemployment Rate	10.67	1.62	7.30	12.70	3,963
1^{st} Income Quartile	0.26	0.11	0.04	0.73	3,963
2^{nd} Income Quartile	0.22	0.09	0.00	0.49	3,963
3^{rd} Income Quartile	0.23	0.06	0.00	0.43	3,963
4^{th} Income Quartile	0.29	0.15	0.02	0.84	3,963

This table reports descriptive statistics on the variables of reading demand and advertising demand as well as on the instrumental variables during the first sample period (only offline world).

In table 2.2, I report the correlation matrix of the main variables for the first sample period (only offline world). There is a positive correlation of 0.3941 between the (natural logarithm of the) number of readers N_{jm} and the (natural logarithm of the) number of advertisement pages A_{jm} in a magazine, suggesting – in line with figure 2.3 – some network effects between readers and advertisers. However, direction and size of these effects remain to be determined in the structural analysis. I also find a negative correlation of -0.6501 between the (natural logarithm of the) number of readers N_{jm} and the copy price cp_{jm}. Finally, I find a positive correlation between the (natural logarithm of the) advertisement price p_{jm} and (the natural logarithm of the) number of advertisement pages A_{jm}, which is *a priori* economically counterintuitive for a demand relationship and therefore indicates the need of instrumental variable techniques to disentangle demand and supply effects in the structural estimation.

Table 2.2: Correlation Matrix: Nested Model (Only Offline World)

	$\ln p_{jm}$	$\ln A_{jm}$	cp_{jm}	$\ln N_{jm}$
$\ln p_{jm}$	1.0000			
$\ln A_{jm}$	0.5014	1.0000		
cp_{jm}	0.0457	0.0154	1.0000	
$\ln N_{jm}$	0.4573	0.3941	-0.6501	1.0000

This table shows the correlation matrix of the main variables of interest for the first sample period (only offline world).

For the second estimation step, the sample size is 45. Table 2.3 provides univariate statistics on this second estimation sample. While the sample size substantially decreases compared to the first sample period, the variables of interest – specifically $\ln p_{jm}$, $\ln A_{jm}$, cp_{jm} and $\ln N_{jm}$ – are remarkably stable on average.[17]

Table 2.4 presents the correlation matrix of the main variables of interest for the second estimation sample. Compared to table 2.2, in this table I also consider the online variables $\ln p_{jw}$, $\ln A_{jw}$ and $\ln N_{jw}$, i.e. price and quantity of online advertising and the number of page impressions. In terms of these additional online variables, table 2.4 shows a positive correlation between the amount of online advertising $\ln A_{jw}$ and the number of page impressions $\ln N_{jw}$, suggesting that there might be inter-group network effects between online advertisers and website viewers, similar as the relationship I find in table 2.2 for the offline market.

It is noteworthy that the correlation coefficients between the offline variables do not change much compared to those from the first sample period shown in table 2.2. In

[17] A t-test on the equivalence of means with unequal variances rejects the null hypothesis of equal means for $\ln p_{jm}$, and cp_{jm} at the 5%-significance level: prices have increased. This cannot be found for logarithmized advertising and reading quantities $\ln A_{jm}$ and $\ln N_{jm}$. However, in absolute terms, A_{jm} and N_{jm} have decreased significantly nevertheless.

Table 2.3: Univariate Statistics: Full Model (Offline and Online World)

	Mean	S.D.	Min	Max	N
PANEL A: ENDOGENOUS VARIABLES: OFFLINE ADS					
$\ln p_{jm}$	10.11	0.44	9.38	11.18	45
$\ln A_{jm}$	5.09	0.75	3.25	6.36	45
PANEL B: ENDOGENOUS VARIABLES: ONLINE ADS					
$\ln p_{jw}$	3.21	0.27	2.30	3.91	45
$\ln A_{jm}$	17.89	1.91	13.59	22.79	45
PANEL C: ENDOGENOUS VARIABLES: OFFLINE READING					
cp_{jm}	3.13	1.44	1.00	7.00	45
$\ln N_{jm}$	14.29	1.12	12.61	16.78	45
PANEL D: ENDOGENOUS VARIABLES: ONLINE READING					
$\ln N_{jw}$	12.71	1.50	9.69	15.88	45
PANEL E: EXOGENOUS VARIABLES: OFFLINE ADS AND OFFLINE READING					
Fashion (%)	10.52	15.91	0.00	49.88	45
Do-it-yourself (%)	8.46	21.45	0.00	84.47	45
Family & Health (%)	8.27	13.04	0.00	65.37	45
Travel & Hobby (%)	23.69	27.56	0.50	85.84	45
Knowledge & Entertainment (%)	39.73	30.42	2.19	95.68	45
Program & Service (%)	9.32	11.42	3.67	83.30	45
ln(# Pages)	6.54	0.59	5.00	7.62	45
PANEL F: EXOGENOUS VARIABLES: ONLINE ADS AND OFFLINE READING					
Redactional content (%)	0.85	0.18	0.29	1.00	45
Content generated by users (%)	0.10	0.15	0.00	0.59	45
Other content (%)	0.05	0.06	0.00	0.32	45
PANEL G: INSTRUMENTAL VARIABLES					
Income: $<$ \$1,000	0.30	0.08	0.14	0.49	45
Income: \$1,000 – \$1,500	0.18	0.02	0.11	0.24	45
Income: \$1,500 – \$2,000	0.17	0.03	0.06	0.23	45
Income: \$2,000 – \$3,000	0.20	0.03	0.11	0.27	45
Income: $>$ \$3,000	0.15	0.06	0.05	0.31	45

This table reports descriptive statistics on the online variables of the full model during the second sample period (offline and online world).

Table 2.4: Correlation Matrix: Full Model (Offline and Online World)

	$\ln p_{jm}$	$\ln A_{jm}$	cp_{jm}	$\ln N_{jm}$	$\ln p_{jw}$	$\ln A_{jw}$	$\ln N_{jw}$
$\ln p_{jm}$	1.0000						
$\ln A_{jm}$	0.4642	1.0000					
cp_{jm}	-0.1932	0.0670	1.0000				
$\ln N_{jm}$	0.7654	0.4850	-0.5685	1.0000			
$\ln p_{jw}$	0.0783	-0.1113	0.4028	-0.1827	1.0000		
$\ln A_{jw}$	0.5020	0.4237	-0.0695	0.5563	0.0264	1.0000	
$\ln N_{jw}$	0.4163	0.3605	-0.0040	0.4977	-0.0290	0.8709	1.0000

This table shows the correlation matrix of the main variables of interest for the second sample period (offline and online world).

particular, the positive correlation between the (natural logarithm of the) number of readers N_{jm} and the (natural logarithm of the) number of advertisement pages A_{jm} in a magazine increases only slightly from 0.3941 in the first sample period to 0.4850 in the second sample period. Similarly, the correlation between the (natural logarithm of the) number of readers N_{jm} and the copy price cp_{jm} changes from -0.6501 to -0.5685. The positive correlation between the (natural logarithm of the) advertisement price p_{jm} and (the natural logarithm of the) number of advertisement pages A_{jm} slightly decreases from 0.5014 to 0.4642.

In sum, these results suggest that neither the offline variables of interest (i.e. $\ln p_{jm}$, $\ln A_{jm}$, cp_{jm} and $\ln N_{jm}$) themselves nor the links between them change much after the rise of the Internet. This supports the assumption that the coefficients estimated in the first step do not change with the rise of the Internet and therefore the two-step estimation methodology outlined above is appropriate.

2.4 Estimation and Results

In this section I estimate the model constructed in section 2.2. As discussed before, estimation proceeds in two steps. I first estimate the nested model with the data from the sample period where online reading and online advertising did not play a significant role. In the second step, these results are plugged into the full model and the resulting coefficients are estimated with data from the year 2009.

2.4.1 Nested Model (Only Offline World)

Recall that the first step consists of the joint estimation of equations (2.7) and (2.10) using GMM with the appropriate instruments described before.

Table 2.5 presents the main results of the first estimation step. Specifications (1) and

Table 2.5: Only-Offline Period: Nested Model

	Without Instruments (Non-linear SUR)				With Instruments (GMM)			
	Without Covariates		With Covariates		Without Covariates		With Covariates	
	(1)		(2)		(3)		(4)	
PANEL A: (INVERSE) DEMAND FOR ADVERTISING								
γ_m	1.280	(142.05)***	1.523	(70.11)***	0.617	(3.80)***	0.220	(1.42)
β_m	0.143	(23.83)***	0.233	(27.33)***	0.403	(7.23)***	0.254	(4.34)***
ln(# Pages)			-0.547	(-14.16)***			1.118	(5.16)***
Fashion (%)			0.001	(1.28)			0.018	(7.06)***
Do-it-yourself (%)			-0.004	(-7.14)***			0.016	(6.06)***
Family & Health (%)			-0.012	(-13.43)***			0.007	(2.60)**
Travel & Hobby (%)			-0.006	(-11.82)***			0.011	(4.36)***
Knowl. & Entertain (%)			-0.005	(-9.95)***			0.004	(2.80)**
Constant	279.121	(13.47)***	6.910	(57.07)***	400.549	(2.15)*	3.198	(7.52)***
PANEL B: READING DEMAND								
ϕ_m	0.641	(37.55)***	-0.066	(-2.40)*	0.580	(8.32)***	0.391	(2.92)**
κ_m	-0.691	(-41.54)***	-0.263	(-18.20)***	-0.761	(-31.94)***	-0.524	(-15.17)***
bi-weekly			-0.686	(-18.20)***			-0.598	(-10.21)***
monthly or less			-1.231	(-20.99)***			-0.853	(-7.40)***
2^{nd} Quarter			-0.059	(-2.18)*			-0.127	(-3.78)***
3^{rd} Quarter			0.008	(0.29)			0.021	(0.70)
4^{th} Quarter			-0.086	(-3.16)**			-0.154	(-4.73)***
ln(# Pages)			0.833	(15.58)***			0.301	(1.41)
Fashion (%)			-0.006	(-8.75)***			-0.008	(-7.32)***
Do-it-yourself (%)			-0.004	(-5.48)***			-0.009	(-5.15)***
Family & Health (%)			-0.005	(-5.41)***			-0.010	(-5.71)***
Travel & Hobby (%)			-0.012	(-18.31)***			-0.014	(-10.32)***
Knowl. & Entertain (%)			-0.007	(-12.27)***			-0.006	(-6.37)***
Constant	12.767	(137.17)***	11.418	(41.44)***	13.243	(40.60)***	13.037	(14.77)***
N	3,963		3,963		3,963		3,963	
P: Hansen's J					0.0000		0.2659	

This table reports the regression results of the first estimation step, i.e. of the nested model during the period when only offline advertisement and offline reading is available. Specifications (1) and (2) employ non-linear Seemingly Unrelated Regression (SUR) while specifications (3) and (4) show the results of General Methods of Moments (GMM) estimation and thus account for endogeneity. Moreover, specifications (1) and (3) do not account for other covariates possibly influencing the endogenous variables whereas specifications (2) and (4) do so. T-statistics are reported in parentheses. ***, **, and * indicate significance at the 0.1%, 1%, and 5% levels, respectively.

(2) show the estimation results of a non-linear Seemingly Unrelated Regression (SUR).[18] In specifications (3) and (4) I use GMM with the instrumental variables described above.

In specification (1), I do not consider any covariates. For advertising demand, I obtain as expected positive and significant network effects from readers to advertisers ($\beta_m = 0.143$). However, the estimation results also predict an upward-sloping demand curve for advertising ($\gamma_m = 1.280$), which is in sharp contrast to my expectations.[19] In terms of reading demand, specification (1) predicts as expected a downward-sloping demand curve ($\kappa_m = -0.691$) and positive and significant network effects from advertisers to readers ($\phi_m = 0.641$), suggesting that readers like advertisements.

When I control for observable characteristics in the magazines – to be specific for the composition of the magazines' contents and the logarithm of the number of content pages – in specification (2), this latter result is reversed: readers dislike advertisements ($\phi_m = -0.066$). Furthermore, compared to specification (1), the price elasticity of reading is decreased in absolute terms ($\kappa_m = -0.263$) and the network effects from readers to advertisers are estimated to be more important than before ($\beta_m = 0.233$). Nevertheless, I still estimate an upward-sloping demand curve ($\gamma_m = 1.523$).

This issue emphasizes the importance to instrument for the potentially endogenous variables in both estimation equations (2.7) and (2.10). I therefore re-estimate both specifications (1) and (2) with GMM using the instruments discussed in sections 2.2.1 and 2.2.2. The corresponding results are shown in specifications (3) and (4) of table 2.5.

In specification (3) I re-estimate the model without covariates. When controlling for endogeneity by using GMM, the positive network effect from readers to advertisers increases ($\beta_m = 0.274$). Additionally, I now also find a downward sloping advertising demand curve ($\gamma_m = 0.617$). In terms of reading demand, the results of specification (3) are very close to those of specification (1). Note that Hansen's J-statistic of overidentifying restrictions is significant even at the 0.1%-level, suggesting that the model is not yet well specified.

When I control for observable characteristics and for endogeneity in specifiction (4), I find – similar to specification (2) – a slightly less elastic reading demand. In comparison to specification (3), γ_m decreases even further.[20] Hansen's J-statistic is insignifcant now (P-value$= 0.2659$), so that the null hypothesis that the model is not overidentified is not

[18]SUR is employed because of common regressors in the equations of advertising demand (2.7) and reading demand (2.10).

[19]Recall that this is also suggested by the positive correlation between $\ln A_{jm}$ and $\ln p_{jm}$ as discussed in section 2.3.2.

[20]In specification (4), γ_m is not significantly different from 0. However, what is important is that γ_m is signifcantly smaller than 1 even at the 0.01% signifcance level (Wald-test).

rejected.

In sum, these results show that advertisers value readers in the traditional German magazine market. This confirms well the corresponding result found in the general empirical literature on two-sided markets that is discussed in the introduction. I also find positive network effects from advertisements to readers, i.e. readers of traditional German magazines like advertisements. Despite the fact that I use a different model for my estimation, this is in line with the findings by Kaiser and Song (2009). Moreover, these findings are encouraging with respect to the overall specification of my model.

2.4.2 Remaining Parameters of the Full Model (Offline and Online World)

I now turn to the estimation of the remaining parameters in the full model. To do so, I plug in the estimates shown in table 2.5 into equations (2.8) and (2.9) and obtain the following system of equations:

$$\begin{aligned}
\ln p_{jm} =& (\hat{\gamma}_m - 1)\ln A_{jm} + \hat{\beta}_m \ln N_{jm} + \frac{1-\rho}{\rho}\ln\left[1 + \frac{X_{jw}\delta_w}{X_{jm}\hat{\delta}_m}\left(\frac{A_{jw}^{\gamma_w}}{A_{jm}^{\hat{\gamma}_m}}\frac{N_{jw}^{\beta_w}}{N_{jm}^{\hat{\beta}_m}}\right)^{\rho}\right] \\
&+ \ln\hat{\gamma}_m + \frac{1}{\rho}X_{jm}\hat{\delta}_m + \nu_{jm} \\
\ln p_{jw} =& (\gamma_w - 1)\ln A_{jw} + \beta_w \ln N_{jw} + \frac{1-\rho}{\rho}\ln\left[1 + \frac{X_{jm}\hat{\delta}_m}{X_{jw}\delta_w}\left(\frac{A_{jm}^{\hat{\gamma}_m}}{A_{jw}^{\gamma_w}}\frac{N_{jm}^{\hat{\beta}_m}}{N_{jw}^{\beta_w}}\right)^{\rho}\right] \\
&+ \ln\gamma_w + \frac{1}{\rho}X_{jw}\delta_w + \nu_{jw} \\
\ln N_{jm} =& \hat{\phi}_m \ln A_{jm} + \phi_w \ln A_{jw} + \psi_w \ln N_{jw} + \hat{\kappa}_m cp_{jm} + Y_{jm}\hat{\theta}_m + \epsilon_{jm} \\
\ln N_{jw} =& \chi_w \ln A_{jw} + \chi_m \ln A_{jm} + \tau_m \ln N_{jm} + Y_{jw}\theta_w + \epsilon_{jw},
\end{aligned} \tag{2.11}$$

where $\hat{\gamma}_m$, $\hat{\beta}_m$, $\hat{\delta}_m$, $\hat{\phi}_m$, $\hat{\kappa}_m$ and $\hat{\theta}_m$ are the estimates of the corresponding coefficients obtained in the first estimation step.

Table 2.6 shows the results of the second estimation step. Similar to table 2.5 from the first step, specifications (1) and (2) present estimation results of non-linear SUR and specifications (3) and (4) those from GMM estimation.

In specification (1), where I do not control for observed characteristics in (inverse) demand for offline and online advertising and in demand for offline and online reading, I find that $\gamma_w = 1.003$. Since this value is not significantly different from 1, it would indicate a flat demand curve for online advertising and therefore it calls for instrumental variable techniques – similar as it was the case for the estimation of offline advertising

Table 2.6: Offline and Online Period: Remaining Parameters

	Without Instruments (Non-linear SUR)				**With Instruments (GMM)**			
	Without Covariates		**With Covariates**		**Without Covariates**		**With Covariates**	
	(1)		(2)		(3)		(4)	
PANEL A: (INVERSE) DEMAND FOR ADVERTISING								
ρ	0.937	(303.16)***	0.671	(66.59)***	0.936	(261.32)***	0.300	(27.80)***
γ_w	1.003	(23.11)***	1.017	(18.69)***	1.041	(17.96)***	0.723	(5.93)***
β_w	-0.024	(-0.46)	-0.023	(-0.31)	-0.064	(-0.81)	0.300	(2.07)*
Red. Content (%)			0.292	(1.00)			0.738	(1.58)
User Content (%)			0.179	(0.59)			0.481	(0.93)
Constant	26.677	(2.87)**	1.879	(4.60)***	21.363	(3.69)***	0.580	(1.30)
PANEL B: READING DEMAND								
ϕ_w	-0.195	(-2.60)**	-0.211	(-4.04)***	-0.120	(-1.74)	-0.617	(-3.72)***
ψ_w	0.313	(2.92)**	0.302	(4.14)***	0.208	(2.13)*	0.890	(3.79)***
χ_w	0.608	(9.10)***	0.637	(10.76)***	0.653	(12.84)***	2.190	(1.43)
χ_m	-0.053	(-0.37)	0.041	(0.27)	-0.006	(-0.08)	0.142	(0.12)
τ_m	0.109	(1.05)	0.060	(0.61)	0.058	(0.69)	1.285	(0.69)
Red. Content (%)			-0.345	(-0.21)			14.295	(0.79)
User Content (%)			-0.817	(-0.42)			6.398	(0.46)
Constant	-0.479	(-0.39)	-0.222	(-0.12)	-0.237	(-0.25)	-32.753	(-1.01)
N	47		45		47		45	
P: Hansen's J					0.1028		0.2647	

This table reports the regression results of the second estimation step, i.e. of the remaining parameters of the full model during the period when both offline and online advertisement and reading is available. Specifications (1) and (2) employ non-linear Seemingly Unrelated Regression (SUR) while specifications (3) and (4) show the results of General Methods of Moments (GMM) estimation and thus account for endogeneity. Moreover, specifications (1) and (3) do not account for other covariates possibly influencing the endogenous variables whereas specifications (2) and (4) do so. T-statistics are reported in parentheses. ***, **, and * indicate significance at the 0.1%, 1%, and 5% levels, respectively.

demand in the first estimation step. Moreover, network effects from online readers to online advertisers are not found to be significant in this specification: β_w is not significantly different from 0. In terms of reading demand, I find that website viewers like advertising ($\chi_w = 0.608$). Most important for the purpose of this study, specification (1) predicts that offline and online advertising are close substitutes ($\rho = 0.937$, which is significantly smaller than 1). When I control for observable heterogeneity in specification (2), most results remain the same. However, ρ decreases significantly to reach a value of 0.671. Similarly, when I use GMM without covariates in specification (3), the results from specification (1) remain almost entirely unchanged.

In specification (4), I control for both observed heterogeneity and endogeneity. In doing so, I find a downward sloping demand curve for online advertising ($\gamma_w = 0.723$) and positive network effects from website viewers to advertisers: advertisers like website surfers. Furthermore, compared to the other specifications, ρ decreases considerably to reach a value of 0.300. Consequently, when one does not control for observed heterogeneity as well as endogeneity, the estimate of ρ is biased upwards. Moreover, the network effect from advertisers to readers χ_w is not estimated to be statistically significant in this specification.

In sum, these results suggest that also in the online market – just as it is the case for the offline market – advertisers value consumers. Moreover, offline and online advertising are found to be substitutes, albeit not perfect ones. I obtain an estimate of ρ which is both far away from $-\infty$ (Leontief specification) and lies between the cases of $\rho = 0$ (Cobb-Douglas specification) and $\rho = 1$ (perfect substitutes specification). The estimate of ρ obtained in the last specification implies a moderate estimated elasticity of substitution between offline and online advertising of $\hat{\sigma} = 1/(1 - \hat{\rho}) = 1.429$, which is much smaller than those obtained from the other three specifications. Indeed, the estimate of ρ will be biased upwards unless the endogenous variables are properly instrumented for as in specification (4).

2.4.3 Discussion

A comment is in order on the behavioral properties of the estimators in terms of bias, consistency and efficiency, and on the implications for the estimates obtained in the second estimation step.

There are three main drivers that may lead to incorrect estimates in the second step. First, if in the first estimation step the model is misspecified, then in the second step inconsistent estimates are plugged into the system of equations (2.6), which in turn yields inconsistent estimates of the remaining coefficients. Second, the assumption that the estimated coefficients from the first step do not change over time might be incorrect. Finally, there may be sample selection in the sense that not all consumer magazines create companion websites and therefore, the substitution process analyzed in this paper is only valid for those which have a companion website.

To shed light on the first two of these issues, it is useful to compare the two-step estimation carried out in this paper with a – hypothetical – one-step estimation of the full model if all necessary data had been available. Suppose that the estimates from the first estimation step are unbiased and that the assumption that the estimates do not change over time is correct. In this case, the estimator in the second step will be consistent and more efficient than the hypothetical one-step estimation. If, by contrast, either the estimates from the first estimation step are biased or the assumption that they do not change over time is wrong, then the estimator in the second step will be inconsistent. However, even in this case the two-step estimator will be more efficient. The reason for the increasing efficiency compared to the one-step estimation is that any additional information – even if it is incorrect – will increase the precision of the estimator.

In order to test robustness of my results with respect to the third issue (i.e. the possibility of sample selection described above), I rerun the regressions of both steps, restricting the sample in the first step to those consumer magazines that create websites and are therefore analyzed in the second step as well. The results of these regressions differ only marginally from those reported in tables 2.5 and 2.6.

2.5 Conclusion

In this paper I have shed light on the interaction between offline and online demand for advertising in the magazines industry, taking into account the two-sidedness of the markets for reading and advertising.

In a first step, I can reproduce the result from previous research that in the traditional two-sided market of consumer magazines, there are positive network effects between readers and advertisers in both directions: advertisers value readers and readers also like advertisements. The reason for this latter result – readers value advertisements – is that when reading a magazine, consumers are always in control because they can easily skip advertisement pages.

Beyond this – and this is the main contribution of my study – offline and online advertising are found to be substitutes albeit not perfect ones. This finding goes in line with the observation made in recent years that advertisers slowly yet increasingly shift their budgets from offline to online advertising. This has potentially important implications for the allocation of advertisement revenue in the future. Essentially, it draws a prospering future for online advertising. As of now, advertisers expect that the positive trend of online advertising will persist in the years to come for three main reasons (see Circle of Online Marketers, 2010). First, for the advertiser costs for online advertising are much more transparent than costs for traditional offline advertising. Second, detailed analyses of the advertising campaigns are much easier for online than for offline advertising. Finally, advertisers can optimize running online advertising campaigns in real time, whereas this possibility is not given in the offline advertising market. As explained in Circle of Online Marketers (2010), these features and possibilities are more important than ever for advertising companies. By sharp contrast, classical magazines probably will have to either cross subsidize their printed editions even more strongly or explore alternative sources of revenue.

3 Estimating Network Effects in Two-Sided Markets without Data on Prices and Quantities

3.1 Introduction

The estimation of network effects in two-sided markets is both particularly important and particularly challenging. It is particularly important because the antitrust economics of two-sided markets differ fundamentally of those of classical one-sided markets (see e.g. Evans, 2003; Wright, 2004) and therefore, the question whether a market is actually two-sided or not is crucial for choosing adequate competition policy measures.[1] The more important the inter-group network effects in a two-sided market are, the bigger are the potential consequences with respect to economic welfare if these network effects are ignored when choosing competition policy measures. Therefore, the inter-group network effects have to be identified and quantified as precisely as possible.

However, the estimation of the network effects paramters in a two-sided market setting is also particularly challenging. Indeed, while the theoretical literature on two-sided markets is constantly growing, only few papers aim to measure the inter-group network effects empirically. This lack of empirical research is largely a consequence of two specific problems arising in the empirical analysis of such markets: the need of complicated estimation procedures and the rather high data requirements.

Both of these problems are due to the fact that the inter-group network effects occur simultaneously and thereby create feedback between the two sides of the market.[2] Various papers have tried to disentangle the two effects of this feedback loop in different two-sided markets. The existing empirical applications include studies on the markets for yellow pages, newspapers/magazines, other media platforms and electronic payment

[1]For a definition and examples of two-sided markets, the reader is referred to section 1.1, and for a more extensive general literature review, see section 1.2.

[2]The analysis can be further complicated if the network effects in one or both directions are negative (see Reisinger, 2004). A typical example is TV advertising, (see Anderson and Coate, 2005; Kind et al., 2005).

systems (in particular payment card networks or Automated Clearing House).

In the simplest case, the existence of inter-group network effects can be identified by a Correlation Analysis. Rysman (2007) analyzes individual network choice and correlations with network reputation, measured by network size and scope. While this method is easy to implement, the direction of the network effect remains unclear, since correlation does not imply causality: The "approach does not allow [...] to distinguish whether consumer usage causes merchant acceptance, merchant acceptance causes consumer usage or both" (Rysman, 2007, p. 27).

Argentesi and Filistrucchi (2007) construct a structural econometric model and estimate it by two-stage least squares in order to derive conclusions on market power in the two-sided market of the newspaper industry. Since the data they have on the two sides does not match in frequency and observation period, they estimate the equations for the two different separately.

Ackerberg and Gowrisankaran (2006) use an Indirect Inference Procedure to estimate causes and magnitudes of network externalities in the Automated Clearing House (ACH) electronic payments system. In a first step, they specify an equilibrium model of interactions between the two sides (banks and customers). Secondly, the structural equilibrium predictions of the model are computed. Finally, these predictions are compared to the actual outcome, i.e. the data.

Rysman (2004) specifies a structural two-sided market model for network competition in the market for yellow pages. Using General Method of Moments (GMM), he estimates simultaneously two demand equations (one on consumer demand for usage and one on advertiser demand for advertising) and the publisher's first-order condition for optimal advertising prices, which is derived from profit-maximizing behavior. The other first order condition is missing since yellow pages are given to readers for free so that prices on the readers' side are constrained to zero. Similarly, Kaiser and Wright (2006) analyze competition in the magazine industry. In their model, two platforms (in this case magazines) compete. Demands on each side is derived from a Hotelling specification and prices are set in a differentiated Bertrand way. They estimate the two demand equations using GMM and are able to derive conclusions on the network externalities and the markups in equilibrium. The advantage of using GMM is that it deals at the same time with the problem of simultaneity arising from the inter-group network effects and the other endogeneity issues. To be specific, simultaneity is addressed by a simultaneous estimation of the demand equations on the two market sides. Endogeneity is addressed by the choice of appropriate instruments.

However, all of the estimation techniques described above put high requirements on the data needed for the analysis. Researchers need data on quantities, prices, covariates as well as appropriate instruments on both market sides. To overcome this problem and in order to use more simple estimation techniques, this paper proposes a semi-structural approach that allows to identify and quantify network effects in two-sided markets absent data on prices and quantities and without using instrumental variable techniques. Instead, only data on equilibrium revenue and on demand and/or cost shifters is required and the resulting regression equations can be easily estimated by Ordinary Least Squares (OLS).

More precisely, I consider a number of profit-maximizing monopoly platforms each of which serves two distinct groups of users that potentially exert inter-group network effects on each other. This monopoly assumption implies that users single-home, i.e. that they use at most one platform and consequently competitive bottlenecks do not exist. One intuitive application of this monopoly setting is the market for regional newspapers in many countries, where each newspaper operates as a monopolist in its respective region. In this example, the two user groups would be readers and advertisers. Network effects between these two groups would be present if advertisers value readers and/or the utility of each reader changes with the number of advertisements in the respective newspaper.

I begin by specifying a structural model of demand and supply. The model is solved for the equilibrium on the two market sides such that endogenous equilibrium prices and quantities are functions of exogenous variables, which may be demand and/or cost shifters.[3] Equilibrium revenue on the two market sides is derived as functions of the shifters. Comparative statics is used to derive conclusions on the inter-group network effects from these revenue equations. Identification is achieved when the reduced-form revenue reacts differently to shocks in demand and/or costs if the two market sides are interconnected rather than if they are not.

The functions are reduced-form revenue equations, since the coefficients to be estimated from them are not the original structural coefficients of the demand and supply equations, but rather combinations of those. This is also the reason why the approach is a semi-structural and not a completely structural one. Revenue results from the optimal prices and quantities in equilibrium and therefore the dependent variable incorporates the overall optimal response to an exogenous change in a shifter through all potentially endogenous variables. As such, the approach follows Panzar and Rosse (1987), who

[3]Intuitively, demand shifters are more promising, since it is difficult to distinguish cost complementarities on the one hand and the difference in the reaction to a shift in costs in one-sided markets and two-sided markets on the other hand.

propose a reduced-form revenue estimation in classical one-sided markets.

By estimating equilibrium revenue instead of demand and supply, I avoid the need of data on prices and quantities as well as of instrumental variables, yet at the cost of losing some economic information which could have been used, had the structural model been directly estimated. Compared to structural models, the drawback of the approach developed in this paper is that after the estimation of the reduced-form revenue equations, not all parameters from the original structural model can be completely identified in all model setups. Specifically, a sufficient test for the existence of network effects is derived when only data on total revenue on the two sides is available. It is not possible to identify size or direction of the network effects in this scenario. If separate revenue data is available, then the test is both necessary and sufficient, and size as well as direction of the network effects are identified.

This paper is structured as follows: In the next section, I propose the semi-structural model of a two-sided market. In section 3.3, the identification of network effects is derived and discussed. Section 3.4 concludes.

3.2 A Semi-Structural Model

In this section, I develop a static, semi-structural model for a two-sided monopoly platform. As discussed above, the main advantage of this approach is that it avoids the need of extensive information on prices, quantities and instrumental variables. Only data on revenues and appropriate demand and supply shifters are needed for the two market sides, and the resulting regression equations can be easily estimated via OLS.

The model follows Armstrong (2006) closely. In particular, prices are not charged per transaction between users from the two sides (see e.g. Rochet and Tirole, 2003, 2006). Rather, each participating user on each side is charged a price for the right to interact on the platform.

Given utility functions of consumers, I derive demand of the two sides of the market. After that I obtain supply from the assumption that the monopolist maximizes joint profit from the two sides, taking into account the demand system derived before. After specifying this structural model, I derive the reduced-form revenue equations from equilibrium prices and number of users, explain how they can be estimated by Ordinary Least Squares and turn to the most important problem of identification in order to draw conclusions on the coefficients of the original structural model, in particular on the existence of network effects. In doing so, I consider two different scenarios, one in which data on total revenue from both market sides is given, but data on revenue from each side is unavailable, and one in which revenue data is available for the two separate sides. The former scenario involves estimation of just one reduced-form revenue equation, whereas in the latter scenario, two equations can be estimated simultaneously.

3.2.1 Utility and Demand

Consider a market with two sides $i = 1, 2$ on each of which there is a measure one of consumers. The utility of a consumer is respectively given by $u_1 = \alpha_1 n_2 - p_1 + \nu_1 + x_1\delta_1$ and $u_2 = \alpha_2 n_1 - p_2 + \nu_2 + x_2\delta_2$, where p_1 and p_2 are the platform's prices for the two groups, n_1 and n_2 are the number of users who are on the other side of the platform and ν_1 and ν_2 represent the consumer type on side 1 and 2, respectively. x_1 and x_2 are demand shifters, for example GDP (if $x_1 = x_2$) or group-specific income (if $x_1 \neq x_2$); their impact is measured by the coefficients δ_1 and δ_2. The economic intuition is that these shifters change the average utility consumers derive from the platform and therefore also the number of consumers who use the platform in equilibrium. The main parameters of interest are α_1 and α_2, which denote the network effects.

Suppose that ν_i is distributed according to the cumulative distribution function $F_i(\nu_i)$ and that u_i is continuous and increasing in ν_i. Without loss of generality ν_i denotes a rank-ordering of the consumer types on side i. Users on the two sides maximize utility, given the expected network size on the other side in equilibrium. Consumers will participate if and only if their utility is non-negative.

Assuming an interior solution, for each side there exists an indifferent consumer of type ν_i^* s.t. $\nu_1^* = p_1 - \alpha_1 n_2 - x_1\delta_1$ and $\nu_2^* = p_2 - \alpha_2 n_1 - x_2\delta_2$.[4] Consumers on side 1 will join if and only if their type ν_1 is bigger than ν_1^*, i.e. if and only if $\nu_1 \geq \nu_1^* = p_1 - \alpha_1 n_2 - x_1\delta_1$. The same holds symmetrically for side 2. The implicit system of demand is then given by $n_1 = 1 - F_1(\nu_1^*) = 1 - F_1(p_1 - \alpha_1 n_2 - x_1\delta_1)$ and $n_2 = 1 - F_2(\nu_2^*) = 1 - F_2(p_2 - \alpha_2 n_1 - x_2\delta_2)$, which can be simplified to:

$$\begin{aligned} n_1 &= 1 - F_1\left(p_1 - \alpha_1 + \alpha_1 F_2\left(p_2 - \alpha_2 n_1 - x_2\delta_2\right) - x_1\delta_1\right) \\ n_2 &= 1 - F_2\left(p_2 - \alpha_2 + \alpha_2 F_1\left(p_1 - \alpha_1 n_2 - x_1\delta_1\right) - x_2\delta_2\right). \end{aligned} \tag{3.1}$$

To close the demand model so that explicit demand can be derived from (3.1), I assume that ν_1 and ν_2 are continuously uniformly distributed in the interval $[0, \nu_{max}]$, i.e. $F_1(\nu_1) = \nu_1/\nu_{max}$ for $0 \leq \nu_1^* \leq \nu_{max}$ and $F_2(\nu_2) = \nu_2/\nu_{max}$ for $0 \leq \nu_2^* \leq \nu_{max}$. This functional assumption of a uniform density of consumers is widely used in the literature,

[4] In this context, the interior solution means that $0 < \nu_1^*\nu_2^* < 1$.

not only in theoretical but also in empirical articles (see e.g. Bresnahan, 1987; Feenstra and Levinsohn, 1995).[5]

For ease of computation, I normalize $\nu_{max} = 1$, so that the distributions of ν_1 and ν_2 are standard uniform ones. The analysis can also be carried out without this additional assumption, but the results become slightly more complex. Under these assumptions, the system of demand (3.1) simplifies to:

$$
\begin{aligned}
n_1 &= \frac{1 - p_1 + x_1\delta_1 + \alpha_1(1 - p_2 + x_2\delta_2)}{1 - \alpha_1\alpha_2} \\
n_2 &= \frac{1 - p_2 + x_2\delta_2 + \alpha_2(1 - p_1 + x_1\delta_1)}{1 - \alpha_1\alpha_2}.
\end{aligned}
\tag{3.2}
$$

3.2.2 Profit Maximization and Supply

Suppose there is a monopoly platform serving demand on the two sides of the market. Consider revenue $R_1 = n_1p_1$ and $R_2 = n_2p_2$ as well as variable costs $C_1(w_1, n_1) = n_1w_1\gamma_1$ and $C_2(w_2, n_2) = n_2w_2\gamma_2$ and fixed costs C_f. w_1 and w_2 are input prices of two cost factors; their impact on variable costs is measured by γ_1 and γ_2, respectively. For simplification, this cost specification assumes that there are constant marginal costs on both sides of the market and that the cost shifters w_1 and w_2 are separable in the market sides, i.e. that factor w_1 does not have any impact on costs on side 2 and, symmetrically, w_2 does not have any impact on costs on side 1.[6]

Other cost specifications are possible. For example, marginal costs need not be constant or there may be cost complementarities between the two market sides. Identification is also possible if variable costs are of the form $C_1(w_1, n_1) = n_1(w_1\gamma_1 + \gamma_1^c)$ and $C_2(w_2, n_2) = n_2(w_2\gamma_2 + \gamma_2^c)$. However, even with the chosen simple cost specification it is difficult to solve for closed-form equilibrium prices and quantities and to derive the original structural parameters from the reduced-form revenue estimators. The more complex the specification of marginal costs is, the more difficult the identification of the original parameters will be.

The monopolist chooses prices for the two sides p_1 and p_2 as to maximize joint platform profit given by:

[5] In principle, other functional assumptions are possible, but it is important that one can solve explicitly for the number of users on each side in equilibrium.

[6] Note that "constant marginal costs" does not mean that the costs may not be shifted as the input factors w_1 and w_2 vary.

$$\begin{aligned}\pi &= \underbrace{(R_1 + R_2)}_{\text{Revenue}} - \underbrace{(C_1(w_1, n_1) + C_2(w_2, n_2) + C_f)}_{\text{Costs}} \\ &= n_1 (p_1 - w_1\gamma_1) + n_2 (p_2 - w_2\gamma_2) - C_f.\end{aligned} \tag{3.3}$$

Substituting (3.2) into profit (3.3) and maximizing with respect to p_1 and p_2 yields optimal prices of the monopolist p_1^* and p_2^*:[7]

$$\begin{aligned}p_1^* =& \frac{2(1 + x_1\delta_1 + w_1\gamma_1) - (\alpha_1 + \alpha_2)(\alpha_1 w_1\gamma_1 + \alpha_2 + \alpha_2 x_1\delta_1)}{4 - (\alpha_1 + \alpha_2)^2} \\ &+ \frac{(\alpha_1 - \alpha_2)(1 + x_2\delta_2 - w_2\gamma_2)}{4 - (\alpha_1 + \alpha_2)^2} \\ p_2^* =& \frac{2(1 + x_2\delta_2 + w_2\gamma_2) - (\alpha_1 + \alpha_2)(\alpha_2 w_2\gamma_2 + \alpha_1 + \alpha_1 x_2\delta_2)}{4 - (\alpha_1 + \alpha_2)^2} \\ &+ \frac{(\alpha_2 - \alpha_1)(1 + x_1\delta_1 - w_1\gamma_1)}{4 - (\alpha_1 + \alpha_2)^2}\end{aligned} \tag{3.4}$$

Substituting (3.4) back into (3.2) yields the equilibrium number of users n_1^* and n_2^*:

$$\begin{aligned}n_1^* &= \frac{2(1 + x_1\delta_1 - w_1\gamma_1) + (\alpha_1 + \alpha_2)(1 + x_2\delta_2 - w_2\gamma_2)}{4 - (\alpha_1 + \alpha_2)^2} \\ n_2^* &= \frac{2(1 + x_2\delta_2 - w_2\gamma_2) + (\alpha_1 + \alpha_2)(1 + x_1\delta_1 - w_1\gamma_1)}{4 - (\alpha_1 + \alpha_2)^2}.\end{aligned} \tag{3.5}$$

In (3.4) and (3.5) all four endogenous variables (n_1^*, n_2^*, p_1^* and p_2^*) are expressed in terms of exogenous variables, their coefficients and the network effects parameters. This is an essential step in the procedure since otherwise the reduced-form revenue equation, which is derived from these two systems of equations, would also consist of potentially endogenous variables on the right-hand side. This in turn would call for instrumental variable estimation techniques.

[7] Suppose that this is a global maximum, which is indeed the case for non-explosive network effects, i.e. $(\alpha_1 + \alpha_2)^2 < 4$ (see e.g. Armstrong, 2006).

3.3 Identfication of Network Effects

Now that the model is solved for equilibrium prices and quantities, the corresponding equations (3.4) and (3.5) are used to express equilibrium revenue as a function of the cost and/or demand shifters and the network effects parameters. Since this involves multiplication of equilibrium prices and quantities, identification of the original structural parameters becomes challenging.

Two different tests are derived in the analysis which follows. First, I show how a sufficient test on network effects can be derived when only data on the sum of the revenues on the two market sides is available. Second, a necessary and sufficient test for network effects can be obtained when revenue data on both sides of the market is available.

3.3.1 A Sufficient Test for Network Effects

Total reduced-form revenue R is given by the sum of the equilibrium revenues from the two market sides:

$$\begin{aligned} R =& R_1 + R_2 = n_1^* p_1^* + n_2^* p_2^* \\ =& \frac{2 + x_1\delta_1(2 + x_1\delta_1) + x_2\delta_2(2 + x_2\delta_2) - (w_1\gamma_1)^2 - (w_2\gamma_2)^2}{4 - (\alpha_1 + \alpha_2)^2} \\ &+ (\alpha_1 + \alpha_2)\frac{(1 + x_1\delta_1)(1 + x_2\delta_2) - w_1\gamma_1 w_2\gamma_2}{4 - (\alpha_1 + \alpha_2)^2}. \end{aligned} \tag{3.6}$$

There are four exogenous variables in (3.6) that shift total equilibrium revenue of the firm, namely x_1 and x_2 as demand shifters and w_1 and w_2 as cost shifters. Rearranging (3.6) yields:

$$R = \lambda_c + \lambda_{x_1}x_1 + \lambda_{x_2}x_2 + \lambda_{x_1^2}x_1^2 + \lambda_{x_2^2}x_2^2 + \lambda_{x_1x_2}x_1x_2 + \lambda_{w_1^2}w_1^2 + \lambda_{w_2^2}w_2^2 + \lambda_{w_1w_2}w_1w_2, \tag{3.7}$$

where the different λ's are combinations of the original structural coefficients: $\lambda_c = \frac{1}{2-(\alpha_1+\alpha_2)}$, $\lambda_{x_1} = \frac{1}{2-(\alpha_1+\alpha_2)}\delta_1$, $\lambda_{x_2} = \frac{1}{2-(\alpha_1+\alpha_2)}\delta_2$, $\lambda_{x_1^2} = \frac{1}{4-(\alpha_1+\alpha_2)^2}\delta_1^2$, $\lambda_{x_2^2} = \frac{1}{4-(\alpha_1+\alpha_2)^2}\delta_2^2$, $\lambda_{x_1x_2} = \frac{-(\alpha_1+\alpha_2)}{4-(\alpha_1+\alpha_2)^2}\delta_1\delta_2$, $\lambda_{w_1^2} = \frac{-1}{4-(\alpha_1+\alpha_2)^2}\gamma_1^2$, $\lambda_{w_2^2} = \frac{-1}{4-(\alpha_1+\alpha_2)^2}\gamma_2^2$, $\lambda_{w_1w_2} = \frac{-(\alpha_1+\alpha_2)}{4-(\alpha_1+\alpha_2)^2}\gamma_1\gamma_2$.

The estimation strategy is to regress total revenue R on the second-order polynomial of the exogenous variables according to equation (3.7), taking into account the corresponding interaction effects. From this regression, the λ's can be consistently estimated by OLS. Solving these λ's for the structural coefficients of the original model yields the following identification:

$$
\begin{aligned}
(\alpha_1 + \alpha_2) &= 2 - \frac{1}{\lambda_c} \\
\delta_1 &= \frac{\lambda_{x_1}}{\lambda_c} \\
\delta_2 &= \frac{\lambda_{x_2}}{\lambda_c} \\
\gamma_1 &= \pm \frac{\sqrt{(1 - 4\lambda_c)\lambda_{w_1^2}}}{\lambda_c} \\
\gamma_2 &= \pm \frac{\sqrt{(1 - 4\lambda_c)\lambda_{w_2^2}}}{\lambda_c}.
\end{aligned}
\tag{3.8}
$$

From (3.8) it can be seen, that the sum $(\alpha_1 + \alpha_2)$ is identified. To be precise, if (3.8) implies $(\alpha_1 + \alpha_2) \neq 0$, then the monopoly platforms under investigation operate in a market with inter-group network effects, i.e. in a two-sided market. According to the first equation of (3.8), testing that $(\alpha_1 + \alpha_2) = 0$ is equivalent to testing that $\lambda_c = 1/2$ in the reduced-form revenue equation (3.7), i.e. the existence of network effects can be tested via the constant of the reduced-form revenue equation.

This special form of the test is due to two specific reasons. The first one is the assumption that $\nu_{max} = 1$, which simplifies the demand functions, thereby also equilibrium prices/quantities and consequently also reduced-form revenue. The second reason is the monopoly assumption, as it implies that the firm can entirely internalize the consumers' reactions to its choice of quantities on both market sides.

However, α_1 and α_2 cannot be identified separately. Economically, this means that both size and direction of the network effects remain indetermined. Even if $(\alpha_1+\alpha_2) = 0$, one cannot exclude the existence of network effects in this scenario, since it is possible that $\alpha_1 = -\alpha_2 \neq 0$, which means that positive and negative network effects cancel each other out. Put differently, this setting provides a sufficient but not necessary test for the existence of network effects.

Demand shock coefficients δ_1 and δ_2 are identified, while for the cost shocks, only the square of the coefficients γ_1 and γ_2 can be identified. Note that not all λ's are needed to derive the solution (3.8). The remaining λ's can therefore be used to construct specification tests for the model.

3.3.2 A Necessary and Sufficient Test for Network Effects

Reduced-form revenue on the two market sides is given by:

$$\begin{aligned} R_i =& n_i^* p_i^* \\ =& \frac{2(1 + x_i\delta_i - w_i\gamma_i) + (\alpha_i + \alpha_j)(1 + x_j\delta_j - w_j\gamma_j)}{4 - (\alpha_i + \alpha_j)^2} \times \\ & \frac{2(1 + x_i\delta_i + w_i\gamma_i) - (\alpha_i + \alpha_j)(\alpha_i w_i \gamma_i + \alpha_j + \alpha_j x_i \delta_i) + (\alpha_i - \alpha_j)(1 + x_j\delta_j - w_j\gamma_j)}{4 - (\alpha_i + \alpha_j)^2} \end{aligned} \tag{3.9}$$

for $i, j = 1, 2$ and $i \neq j$. As before, there are four exogenous variables in (3.9) that shift equilibrium revenues of the firm on both sides of the market: x_1, x_2, w_1 and w_2. Rearranging (3.9) yields:

$$\begin{aligned} R_i =& \lambda_c^{(i)} + \lambda_{x_1}^{(i)} x_1 + \lambda_{x_2}^{(i)} x_2 + \lambda_{x_1^2}^{(i)} x_1^2 + \lambda_{x_2^2}^{(i)} x_2^2 + \lambda_{x_1 x_2}^{(i)} x_1 x_2 \\ & + \lambda_{w_1}^{(i)} w_1 + \lambda_{w_2}^{(i)} w_2 + \lambda_{w_1^2}^{(i)} w_1^2 + \lambda_{w_2^2}^{(i)} w_2^2 + \lambda_{w_1 w_2}^{(i)} w_1 w_2 \\ & + \lambda_{x_1 w_1}^{(i)} x_1 w_1 + \lambda_{x_2 w_2}^{(i)} x_2 w_2 + \lambda_{x_1 w_2}^{(i)} x_1 w_2 + \lambda_{x_2 w_1}^{(i)} x_2 w_1 \end{aligned} \tag{3.10}$$

for $i = 1, 2$ and where the $\lambda^{(i)}$'s are combinations of the original structural coefficients.

Estimation is to regress revenues of the two market sides R_1 and R_2 separately on the second-order polynomial in the exogenous variables, taking into account their interaction effects, as indicated in (3.10).[8] The resulting estimates of the $\lambda^{(1)}$'s and the $\lambda^{(2)}$'s are consistent and can be used to identify the structural parameters of the original model as follows:

[8] In principle, the two reduced-form revenue equations of system (3.10) could also be estimated jointly with Seemingly Unrelated Regression (SUR) estimation. However, as the explanatory variables are identical in both regression equations, there are no efficiency gains from using SUR rather than OLS (see Judge et al., 1988, p. 452).

$$
\begin{aligned}
\alpha_1 &= 1 - \frac{\lambda_c^{(2)}}{\left(\lambda_c^{(1)} + \lambda_c^{(2)}\right)^2} \\
\alpha_2 &= 1 - \frac{\lambda_c^{(1)}}{\left(\lambda_c^{(1)} + \lambda_c^{(2)}\right)^2} \\
\delta_1 &= \frac{4\left(\lambda_c^{(1)} + \lambda_c^{(2)}\right) - 1}{5\left(\lambda_c^{(1)}\right)^2 - \lambda_c^{(1)} + 6\lambda_c^{(1)}\lambda_c^{(2)} + \left(\lambda_c^{(2)}\right)^2}\lambda_{x_1}^{(1)} \\
\delta_2 &= \frac{4\left(\lambda_c^{(1)} + \lambda_c^{(2)}\right) - 1}{5\left(\lambda_c^{(2)}\right)^2 - \lambda_c^{(2)} + 6\lambda_c^{(1)}\lambda_c^{(2)} + \left(\lambda_c^{(1)}\right)^2}\lambda_{x_2}^{(2)} \\
\gamma_1 &= \frac{1 - 4\left(\lambda_c^{(1)} + \lambda_c^{(2)}\right)}{\left(\lambda_c^{(1)} - \lambda_c^{(2)}\right)^2 + \lambda_c^{(2)}\left(1 - 4\lambda_c^{(2)}\right)}\lambda_{w_1}^{(1)} \\
\gamma_2 &= \frac{1 - 4\left(\lambda_c^{(1)} + \lambda_c^{(2)}\right)}{\left(\lambda_c^{(1)} - \lambda_c^{(2)}\right)^2 + \lambda_c^{(1)}\left(1 - 4\lambda_c^{(1)}\right)}\lambda_{w_2}^{(2)}.
\end{aligned}
\tag{3.11}
$$

(3.11) shows that if revenue on both sides is given, then α_1 and α_2 are identified. Economically, this means that both size and direction of the network effects parameters are identified. Moreover, the test on network effects in this scenario is both necessary and sufficient.

As before, conclusions on the structural coefficients of network effects α_1 and α_2 can be made using only the intercepts of the reduced-form regression equations. In particular, summing up the first two equations of (3.11) and noting that $R_1 + R_2 = R$ implies $\lambda_c^{(1)} + \lambda_c^{(2)} = \lambda_c$, one reobtains the sufficient test discussed in section 3.3.1 for the scenario when only total revenue is available. Identification of α_1 and α_2 is now achieved because in contrast to this former scenario, separate revenue constants $\lambda_c^{(1)}$ and $\lambda_c^{(2)}$ are available for the two revenues R_1 and R_2 and their individual contribution to the total revenue constant λ_c can be determined. The special form is again due to the simplifying assumption $\nu_{max} = 1$ and the monopoly setting.

3.4 Conclusion

In this paper, I propose necessary and sufficient tests for network effects in two-sided markets absent data on prices and quantities. This is possible because in such markets, cost and/or demand shocks have a different influence on equilibrium revenue on the two sides than they have in markets which are not interconnected.

When only data on the sum of the revenues on the two sides is available, it is possible to derive conclusions on the existence of network effects; however, the test is only a sufficient one and it is impossible to identify the direction of the network effects. If, by contrast, revenues on both sides are available, the test is a necessary and sufficient one and the direction of network effects is identified.

There are several limitations in and potential extensions of the approach presented in this paper. First, the model is only valid for monopoly platforms. It may be an interesting extension to construct a similar model for duopolies, oligopolies and perfect competition. Second, in order to be able to solve for the equilibrium prices and quantities in the monopoly model of a two-sided platform, I have to assume a uniform (or a different simple) distribution of the consumers' heterogeneity parameters ν_1 and ν_2 on both sides of the market. In reality, consumers' heterogeneity may be different. The conclusions of the semi-structural approach proposed in this paper rely on this specifc structural functional assumption. However, the model itself provides appropriate specification tests for this underlying assumption. Therefore, the approach can be used as a good starting point for researchers who aim to identify whether a platform is two-sided or not.

4 Skin in the Game: Evidence from the Online Social Lending Market

4.1 Introduction

The functioning of markets crucially depends on the matching of demand and supply, and this holds in particular for financial markets such as the lending market. Borrowers and lenders face substantial information asymmetries, which may eventually lead to the breakdown of this market as described by Akerlof (1970) and Stiglitz and Weiss (1981) – and as observed in the recent financial crisis. Banks have traditionally taken the role of financial intermediaries to screen and monitor potential borrowers by using public and private information to overcome – at least partly – these information asymmetries and to allow the lending market to work, i.e. to give creditworthy borrowers access to credit at sustainable interest rates that incorporate the borrowers' risk of default. Their commitment to the scrutiny of screening and monitoring and thus the forbearance from unscrupulous lending to informationally disadvantaged borrowers such as retail customers has traditionally been secured by their skin in the game, as described in Holmström (1979) and Holmström and Tirole (1997). However, the widespread use of loan securitization and the originate-to-distribute model have altered the incentives for financial intermediaries and raised the important question whether and to what extent the lack of skin in the game has affected the quality of lending decisions. Discussion about this question has been at the forefront of the regulatory and academic debate about the financial crisis.[1] Further, with the recent advances in information technology, new lending platforms have emerged that do not rely on the existence of a financial intermediary any more and in which lenders and borrowers do not have the chance for personal interaction, as for example described in Ravina (2008). Important open questions are how markets can responsibly match demand and supply despite the lack

[1] For example, President Obama motivated the creation of the Consumer Financial Protection Agency as follows (see The White House, 2009): "Millions of Americans who have worked hard and behaved responsibly have seen their life dreams eroded by the irresponsibility of others and the failure of their government to provide adequate oversight. Our entire economy has been undermined by that failure."

of a financial intermediary and skin in the game as well as which conditions have to be fulfilled and what incentives have to be given to market participants to protect retail customers from unscrupulous lending. While these questions have relevance for many financial markets, the lack of data makes it often very difficult to find clear evidence.

We thus address some of these questions by examining a clearly defined major change on the online social lending platform Prosper.com, on which lenders can give their money directly to borrowers without the intermediation of a financial institution. Prosper.com has attracted over 385,000 requests for loans with a total volume of more than $2,800,000,000 since its inception in 2006. As an outcome, 36,268 of these loan requests have resulted in actual loans with a volume of $211,000,000. Prosper.com has thus developed into the market leader for online social lending and can be seen as an ideal and clean opportunity for our analysis as it provides on its webpage detailed information on individual borrowers, their loan requests, funding success, interest rates, and subsequent loan performance.

We are able to examine which incentives work well in this market as well as identify mechanisms that lead to a deterioration in lending quality, using a difference-in-difference methodology and analyzing the effects of a major change in the way the lending platform operates. One important mechanism in this market is the creation of self-organized groups that are headed by a group leader and joined voluntarily by further members. The group leader is allowed to grant or deny members access to their group, ask for verification of the information provided by the group members and define the purpose of the group as well as the nature and interests of its members. In particular, the group leader can endorse and submit bids for the borrower listings in her group, i.e. put her money where her mouth is, or have "skin in the game". Groups can have the equivalent of an origination fee wherein the group leader is allowed to charge a fee for his role in matching demand and supply for loans. This fee regularly comprises an immediate closing fee and additional interest over the lifetime of the loan. Prosper.com abolishes this group leader reward on 09/12/2007, following an announcement on 09/05/2007. This imposed change on the group leader provides us with a unique opportunity to analyze the functioning of the market before and after this change in the reward structure for the group leader. Importantly, we can see the behavior of the same group leaders and groups before and after the removal of origination fees and assess differences in the kinds of loans originated and their performance.

We find that group rewards have an adverse effect, as we document remarkable differences for lending outcomes and in particular default rates before and after the change of the reward structure. When group leaders can still earn rewards for successful listings in their groups, the default rates are substantially higher for loans with than for loans without group leader bids and endorsements. From an economic standpoint, it still pays for the group leader to endorse or submit bids even for weaker listings. The successful closure of these listings provides him with a reward that exceeds the losses from the increased likelihood of default, while other lenders and borrowers lose on these loans. In strict contrast, after the change in the reward structure when the group leader does not receive any fees for a successful closure of a listing any more, group leader bids and endorsements are used much more responsibly and are thus associated with significantly lower borrower default rates.

Similarly, even before the elimination of group leader rewards, a group leader bid and endorsement is credible when the group leader contributes a substantial fraction to the requested loan amount. In this case, the default rates are significantly lower than for other loans and almost identical to those for loans after the elimination of group leader rewards. These results suggest that a group leader has the right incentives to screen only if he has substantial skin in the game and is severely hurt by losing money when a borrower defaults. This evidence has important implications for the current debate about the proper protection of retail customers in financial markets. In particular, it suggests that only originators who retain a substantial share of the originated loan have the right incentives to screen loans efficiently and make responsible lending decisions that do not hurt borrowers and co-lenders.

Our paper is related to different strands of the literature. First, it deals with the general questions raised in Akerlof (1970) and Stiglitz and Weiss (1981) of how to match demand and supply and thus enable the lending market to work. We provide evidence how group leader bids and endorsements as well as group leaders' skin in the game provide credible signals to other lenders and thus induce them to bid on these listings. The paper thus directly relates to the literature that focuses on the unobservable actions by the lender in checking potential borrowers' creditworthiness. The theoretical work by Holmström (1979) and Holmström and Tirole (1997) as well as the empirical work by Sufi (2007) stress the importance of the share of the loan retained by financial intermediaries to overcome information asymmetries. Second, our paper relates to the growing literature on irresponsible advice and lending by financial intermediaries and the resulting need for regulatory intervention and consumer protection, such as for example Bolton et al. (2007), Bergstresser et al. (2009), and Inderst and Ottaviani (2009). Third, we analyze which particular role important concepts from the banking literature play in this context. One important related concept is the differentiation between hard and soft information such as in Stein (2002), and Berger et al. (2005). An important change due to the use of new technologies in finance such as online lending is a greater reliance on hard relative to soft information in financial transactions. At the same time, information technology may lead to the hardening of soft information, i.e. the possibility to transform the nature of the information from soft into hard as for example in credit ratings. Another important related concept is the inherent risk of free-riding in monitoring when a larger number of lenders face a single borrower, along the lines in Bolton and Scharfstein (1996). Finally, there is a growing number of papers that analyze the lending behavior on Prosper.com. Hulme and Wright (2006) provide an overview of the historical origins and contemporary social trends of online social lending and conduct a case study of the world's first online social lending platform, Zopa. Ravina (2008) and Pope and Sydnor (ming) analyze whether there is discrimination on Prosper.com in terms of socio-demographic variables such as race and gender. These characteristics are taken care of by the difference-in-difference methodology employed in this paper, assuming their distribution is time-invariant across the different groups. Iyer et al. (2009) test whether lenders can infer soft information in Prosper. Lin et al. (2009) test which role social networks and in particular "the company that borrowers keep", i.e. the borrowers' friends, play for the lending outcome. In our study, we focus on group leader bids and endorsements as mechanisms used by the group leader to promote listings, and we specifically examine the consequences of the elimination of group leader rewards for funding success, the resulting interest rate, and loan performance. This helps us to better understand the implications of the use of different incentives in consumer lending in this market and in particular the importance of skin in the game.

The rest of the paper is structured as follows. The next section describes the institutional setting on the platform and provides an overview over the data. Section 4.3 presents the analysis and the univariate and multivariate results. Section 4.4 concludes.

4.2 Institutional Setting and Data

4.2.1 The General Setup

Prosper.com provides a basis for the interaction between two sides: on the one side the potential borrowers, who are looking for money for some specific purpose; on the other side the potential lenders, who are interested in opportunities and projects to invest their money into.[2] After registering on the platform, borrowers can post a listing in which they ask for money and provide different types of information so that potential lenders

[2]Institutions are not allowed on Prosper.com during the sample period, so only private persons may serve as borrowers or lenders.

can better assess their creditworthiness. These types of information can be classified into hard and soft information:

- *Hard information*
  - *On the borrower*: Prosper.com assigns a unique identification number to each borrower and requires him to provide his social security number, driver's license number, and bank account information so that Prosper.com can verify his identity and obtain his Experian Scorex PLUS^{SM} credit report. Of particular importance here is the credit grade, which ranges from AA for the best customers over A, B, C, D, and E to HR for the worst customers and which is assigned to potential borrowers based on their Experian credit score. The credit report, which is not reviewed or verified by Prosper.com, also includes the borrower's default history, which is thus observable by potential lenders.
  - *On the listing*: Borrowers set the amount they request, which is between \$1,000 and \$25,000, as well as the maximum interest rate they are willing to pay. In some states, there are interest rate caps, while in the other states the maximum interest rate may go up to 35% – an interest rate cap set by Prosper.com.
- *Soft information*

  This information is provided by the borrower herself and only some of it is verified. Examples of this soft information are borrower state, income range, and house ownership. Additionally, the borrower has the possibility to post one or more photos, e.g. of her or the object that she wants to finance with the loan. Borrowers can explain what they want to spend the money on, how they intend to pay it back by providing a budget, and why they are particularly reliable and trustworthy.

Lenders have the possibility to screen the listings and can place one or several bids of at least \$50 on any of them at any interest rate below or equal to the maximum interest rate requested by the borrower. These bids cannot be canceled or withdrawn. The bidding on the listing is performed as an open uniform-price auction in which everybody can observe each other's actions. As long as the aggregate supply on a listing does not exceed the borrower's demand, bidders can see the amount of the other bids, but not the interest rates of those bids. They only observe the maximum interest rate that the borrower is willing to pay. Once the aggregate supply exceeds the borrower's demand, bidders can also see the marginal interest rate so that they know which rate they have to underbid to be able to serve as a lender. As a consequence, lenders who offer the highest interest rates are outbid, so that the resulting interest rate is bid down until the duration of the listing expires and the listing becomes a loan. Alternatively, borrowers can also choose that the listing is closed and the loan is funded as soon as the total amount bid reaches the amount requested. In the end, all winning bidders receive the same interest rate, which is the marginal interest rate. In case the total amount bid does not reach or exceed the amount requested within the duration time, the listing expires and no transaction takes place. All loans on Prosper.com are 36-months annuity loans, which can be paid back in advance though. The platform makes money from charging fees to borrowers and lenders once a listing is completely funded and becomes a loan. Borrowers pay – depending on their credit grade – a one-time fee (between 1% and 5% of the loan amount), which is subtracted from the gross loan amount. Lenders pay a 1% annual servicing fee.

A borrower who defaults on his loan is reported to credit bureaus so that this information is recorded in the borrower's credit report. Prosper.com uses collection agencies to recover the outstanding balances, and the fees for these agencies are borne by the defaulting borrowers' lenders. Loans are unsecured and there is no second market for these loans unless they become overdue; Prosper.com then reserves the right to sell the loans to outside debt buyers.

On Prosper.com, platform members can organize themselves in groups in order to facilitate the process of borrowing and lending as well as the interaction between each other. Each user can form a group by defining the purpose of the group as well as the nature and interests of its members and thus become a group leader. Each user can be member (and thus group leader) of at most one group. The group leader administers her group and can additionally act as a lender and / or borrower on the platform. Furthermore, the group leader has the right to grant or deny other users access to her group and ask for verification of the information that these users provide. Many group leaders request additional information from potential borrowers, and this process is referred to as "Vetting". Furthermore, some group leaders request to review every listing before it is posted in the group. Finally, there are group leaders who explicitly offer help to the potential borrower in writing and designing the listing.

On Prosper.com, platform members can organize themselves in groups in order to facilitate the process of borrowing and lending as well as the interaction between each other. Each user can form a group by defining the purpose of the group as well as the nature and interests of its members and thus become a group leader. Each user can be member (and thus group leader) of at most one group. The group leader administers her group and can additionally act as a lender and / or borrower on the platform.

Furthermore, the group leader has the right to grant or deny other users access to her group and ask for verification of the information that these users provide. Many group leaders request additional information from potential borrowers, and this process is referred to as "Vetting". Furthermore, some group leaders request to review every listing before it is posted in the group. Finally, there are group leaders who explicitly offer help to the potential borrower in writing and designing the listing.

The group leader can exploit this potential informational advantage and the fact that everybody can observe each other's actions to promote in different ways the listings posted in her group among potential lenders: she can place a bid on the respective listing, thereby potentially signaling a financial commitment to the trustworthiness of the borrower. Furthermore, the group leader can write an endorsement for the potential borrower, i.e. a short text in which she describes why this respective borrower is particularly trustworthy. While bids and endorsements can also be made by other members of Prosper.com, we concentrate on the analysis of bids and endorsements by the informationally advantaged group leaders, who are also much more active than other group members and are the key facilitators in their respective groups. Group leader bids and group leader endorsements are often given together. We thus use the following approach. First, in the univariate analysis, we consider the two signaling mechanisms separately. Later, in the multivariate analysis, we analyze group leader bids and group leader endorsements simultaneously.

4.2.2 Reward Groups, No-Reward Groups, and the Elimination of Group Leader Rewards

Apart from the fact that groups aim at different purposes and people, they are very heterogeneous by nature: Group leaders may either provide their service for free, for example because of the interest they can earn on the loans to which they lend money or simply the benefits from social interaction or prestige, or charge a fee on loans closed in their group.[3] Therefore, in our analysis we distinguish between *no-reward groups* and *reward groups*. More precisely, we define a group as a reward group if the group leader requires a group leader reward at least for one listing in her group. Otherwise, the group is defined as a no-reward group.

Prosper.com started its business officially in 2006. Since then, there have been several

[3] The group leader obtains a one-time reward ("match reward", 0.5% of the loan amount except for E-loans and HR-loans) once the listing is completely funded and a monthly payment ("payment reward", 1% p.a. for AA-loans and A-loans, 2% p.a. for B-loans, C-loans and D-loans, 4% p.a. for E-loans and HR-loans.). Alternatively, the group leader can also choose to only partly capture this reward.

policy changes on the platform to adjust the business model to changes in the macroeconomic environment and to the constantly better understanding of how online social lending works. Figure 4.3 in the appendix provides a corresponding timeline of these policy changes. In our analysis, we focus on one specific policy change: the elimination of group leader rewards, which takes place on 09/12/2007. Prosper.com motivates the elimination of group leader rewards in its announcement by "(t)he original philosophy [...] to enable borrowers in close-knit communities to leverage the reputation and peer pressure of their group [...], where compensation is not the dominant motivation for the group leader's services." This event constitutes an imposed change on leaders of reward groups and systematically changes their incentives in the loan granting process. It thus represents an ideal event to analyze how group leaders react to a sudden change in incentives. To exclude possible influences of other significant policy changes, we restrict our analysis to the loans originated between 02/13/2007 and 04/15/2008 in which no other significant policy change occurs and follow their performance until 03/01/2010.[4] On 02/12/2007, Prosper.com redefines the credit grades E and HR, excludes borrowers without any credit grade from the platform, changes the borrower closing fee from 1% to 2% for the credit grades E and HR and the lender servicing fee from 0.5% to 1% for the credit grades B-HR. Also, endorsements for friends are introduced in addition to group leader endorsements. On 04/15/2008, Prosper.com increases the lender servicing fee for AA-loans from 0% to 1%. The policy change of interest in our study – the elimination of group leader rewards – is thus well centered in the sample period.

4.2.3 Descriptive Statistics

Until today, 36,268 loans have been originated out of more than 385,000 listings on Prosper.com. The total amount funded exceeds $211,000,000. The company makes a snapshot of its entire public data available on its website for download and data analysis. After restricting the sample period as discussed above, we obtain a final sample of 153,541 listings, 34,858 of which are posted in groups.

[4]During the sample period, there are two minor policy changes: On 10/30/2007, Prosper.com changes the lender servicing fee from 0.5% to 1% for A-loans and from 0.5% to 0% for AA-loans. Moreover, from this date on Prosper.com allows borrowers who already have a current loan to create a new listing in order to obtain a second loan. Second loans are allowed only for borrowers whose first loan has been active for some time and whose two loans together do not exceed the maximum amount of $25,000. To control for this latter policy change, we remove from the analysis the corresponding listings in which borrowers apply for second loans. On 01/04/2008, Prosper.com changes the borrower closing fees from 1% to 2% for the credit grades A and B, from 1% to 3% for the credit grades C and D, and from 2% to 3% for the credit grades E and HR.

Table 4.1 provides the summary statistics for the most important variables.[5] Panel A shows the distribution of listings by credit grades and by groups. Most listings are either posted outside a group (118,683) or in a reward group (32,966); much fewer listings are posted in no-reward groups (1,892). Listings with the credit grade HR present by far the most dominant group of listings with 66,734 observations, again mostly outside a group and in reward groups.

Table 4.1: Summary Statistics

	No Group	**No-Reward Groups**	**Reward Groups**	**Overall**
PANEL A: DISTRIBUTION OF LISTINGS				
AA / A	7,641	301	1,641	**9,583**
B	6,532	146	1,839	**8,517**
C	12,572	293	3,648	**16,513**
D	18,896	346	5,529	**24,771**
E	21,005	261	6,157	**27,423**
HR	52,037	545	14,152	**66,734**
Total Number of Listings	118,683	1,892	32,966	**153,541**
PANEL B: DISTRIBUTION OF LOANS				
AA / A	2,303	181	659	**3,143**
B	1,366	73	540	**1,979**
C	1,572	119	839	**2,530**
D	1,258	130	904	**2,292**
E	514	63	495	**1,072**
HR	432	88	647	**1,167**
Total Number of Loans	7,445	654	4,084	**12,183**
PANEL C: GROUP-SPECIFIC INFORMATION				
Group Leader Bid		45.8%	32.0%	**32.7%**
Group Leader Endorsement		32.8%	12.4%	**13.5%**
Vetting		28.6%	9.4%	**10.4%**
Listing Review Requirement		66.0%	40.7%	**42.1%**
Group Leader Offers Help		18.1%	7.8%	**8.3%**

In this table we report – by group type – summary statistics on the most important variables. Panel A shows the distribution of listings (i.e. of requests for borrowing money) by the different credit grades from AA/A (best) to HR (worst). Panel B shows the corresponding distribution of loans (i.e. of successfully and completely funded requests for borrowing money). Panel C reports general group-specific shares, in particular the share of listings with at least one group leader bid and the share of listings with a group leader endorsement. “Vetting” denotes that the group leader claims to review information sent by the borrower (e.g. diploma or certificates). “Listing Review Requirement” denotes that the group leader checks the listing before it is opened for bidding by potential lenders. “Group Leader Offers Help” denotes that the group leader offers to support the borrower in writing and designing the listing.

From panel B of Table 4.1 we see that this does not hold true for the distribution of loans. From the 12,183 loans, only 1,167 originate from successfully funded HR-listings, while there are by far more AA/A-loans (3,143). Only for E-loans, the number of loans is

[5] Variable definitions for all variables in the tables of the paper are given in Table 4.8.

smaller than for HR-loans. The results in panel B also suggest that the listing probability is highest in no-reward groups, followed by that in reward groups and outside groups. The number of loans in no-reward groups of 654 constitutes almost 35% of the number of listings of 1,892 in these groups, while this rate decreases to about 12% for reward groups and 6% outside groups.

In panel C of Table 4.1, the information on group-specific characteristics is summarized. Despite the fact that they are not compensated for their work, group leaders are relatively more active in no-reward groups than in reward groups in terms of bidding and endorsing listings. They are also more involved in terms of vetting, i.e. they review and certify the information given to them by the potential borrowers, reviewing listings, and offering help to the borrower. For example, the share of listings with at least one group leader bid is considerably higher in no-reward groups (45.8%) than in reward groups (32.0%).

4.3 Empirical Analysis and Results

4.3.1 Univariate Analysis

Group Leader Bids and Group Leader Endorsements

Group leaders can use bids and endorsements as two important mechanisms to promote listings in their groups. However, the existence of rewards for group leaders may create adverse incentives for these group leaders. Rewards for successful listings may induce them to use bids and endorsements to persuade other lenders to bid even on weak listings, by making other lenders believe that these listings are creditworthy. Thus, in the first step, it is important to understand how bids and endorsements are used in no-reward and reward groups and which outcomes are associated with them. In the observed period, group leaders bid on 32.7% of the listings and these bids tend to be successful: among all first group leader bids on a listing, only 13% are outbid. Mostly, these bids constitute small amounts – very often \$50 or \$100 – so that the median amount of the first group leader bid is \$70. Usually, these bids are placed very fast. Indeed, if a group leader bids, her first bid is typically also the first overall bid on the respective listing.

Table 4.2 analyzes for no-reward and reward groups the listing success, interest rates, and loan performance based on whether the group leaders bids on or endorses a listing or whether he abstains from either of the two.

Panel A of Table 4.2 shows how success rates of listings are related to group leader bids and group leader endorsements. In no-reward groups, success rates for listings with

Table 4.2: Listing Success, Interest Rates, and Loan Performance by Listing Promotion Mechanism (Group Leader Bids and Group Leader Endorsements)

Panel A: Listing Success

	No-Reward Groups (N=1,892)					Reward Groups (N=32,966)				
Credit Grade	**None**	**With GL-Bid**		**With GL-Endorsement**		**None**	**With GL-Bid**		**With GL-Endorsement**	
AA/A	39.5%	78.1%	(-6.81)***	81.9%	(-7.54)***	31.2%	50.0%	(-7.56)***	69.6%	(-13.29)***
B	34.3%	64.6%	(-3.67)***	76.6%	(-5.00)***	20.8%	38.5%	(-8.06)***	60.7%	(-13.87)***
C	21.3%	60.6%	(-7.31)***	70.8%	(-8.31)***	14.9%	33.2%	(-12.21)***	54.0%	(-17.61)***
D	13.2%	56.0%	(-9.37)***	68.9%	(-10.97)***	9.7%	26.4%	(-15.04)***	45.5%	(-19.43)***
E	9.5%	42.5%	(-6.22)***	55.4%	(-7.25)***	3.4%	18.0%	(-15.49)***	31.2%	(-15.28)***
HR	4.3%	32.4%	(-8.38)***	33.1%	(-7.58)***	2.0%	11.1%	(-17.54)***	19.6%	(-16.03)***
Total	**16.6%**	**52.8%**	**(-17.22)*****	**60.6%**	**(-18.97)*****	**6.9%**	**22.4%**	**(-35.17)*****	**39.3%**	**(-41.37)*****

Panel B: Interest Rates

	No-Reward Groups (N=654)					Reward Groups (N=4,084)				
Credit Grade	**None**	**With GL-Bid**		**With GL-Endorsement**		**None**	**With GL-Bid**		**With GL-Endorsement**	
AA/A	9.3%	9.3%	(-0.11)	9.5%	(-0.37)	11.0%	11.4%	(-2.10)**	11.7%	(-2.79)***
B	13.4%	12.4%	(1.34)	12.9%	(0.61)	15.2%	14.6%	(1.65)*	14.9%	(0.85)
C	15.8%	15.6%	(0.22)	15.6%	(0.17)	18.2%	16.8%	(4.73)***	17.1%	(3.49)***
D	19.2%	17.4%	(1.94)*	17.1%	(2.10)**	20.9%	19.7%	(3.97)***	19.6%	(4.22)***
E	21.5%	20.6%	(0.62)	20.4%	(0.72)	24.8%	23.8%	(2.24)**	23.5%	(2.58)***
HR	24.7%	19.7%	(2.37)**	20.7%	(1.89)*	26.1%	24.2%	(4.50)***	24.3%	(4.06)***
Total	**14.8%**	**15.5%**	**(-1.20)**	**15.4%**	**(-1.03)**	**18.7%**	**18.8%**	**(-0.53)**	**18.5%**	**(0.77)**

Panel C: Loan Performance

	No-Reward Groups (N=654)					Reward Groups (N=4,084)				
Credit Grade	**None**	**With GL-Bid**		**With GL-Endorsement**		**None**	**With GL-Bid**		**With GL-Endorsement**	
AA/A	2.8	6.3	(7.70)***	4.5	(3.97)**	6.6	10.6	(14.16)***	11.0	(14.26)***
B	7.7	3.5	(-5.54)***	7.0	(-0.81)	13.3	15.8	(6.42)***	15.5	(5.27)***
C	8.8	10.3	(2.04)**	8.7	(-0.09)	16.7	16.8	(0.34)	16.3	(-1.21)
D	9.6	10.5	(1.02)	9.5	(-0.13)	16.8	17.5	(2.05)**	16.9	(0.21)
E	19.4	13.2	(-4.33)***	12.4	(-4.79)***	18.5	22.9	(9.21)***	25.5	(12.89)***
HR	31.4	21.1	(-5.66)***	22.9	(-4.62)***	23.7	26.4	(5.70)***	29.1	(10.26)***
Total	**10.6**	**11.4**	**(2.10)****	**10.9**	**(0.87)**	**15.7**	**18.9**	**(20.79)*****	**19.0**	**(19.98)*****

In this table we report univariate results by listing promotion mechanism (group leader bids / group leader endorsements) and credit grade. The table distinguishes between No-Reward Groups and Reward Groups. Panel A shows success rates of listings (i.e. of the requests for borrowing money) by the different credit grades from AA/A (best) to HR (worst). Panel B shows the corresponding interest rates of loans (i.e. of the successfully and completely funded requests for borrowing money). Panel C shows failure rates of loans (per 1,000 loan-days). In this panel, any payment which is not made on time is considered as a failure, so that failure events are late payments, charge-offs and defaults. T-statistics of the test on equality between "With GL-Bid" and "None" as well as between "With GL-Endorsement" and "None" are reported in parentheses for both No-Reward Groups and Reward Groups. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

a group leader bid (52.8%) or a group leader endorsement (60.6%) are much higher than for those which have neither (16.6%). This is true for all credit grades, which shows that both group leader bids and group leader endorsements increase the probability of funding regardless of the riskiness of the listing. The analysis of reward groups draws a similar picture: here, only 6.9% of the listings without a group leader bid and without a group leader endorsement are funded, while the listing success is significantly increased by group leader bids (22.4%) and group leader endorsements (39.3%).

From panel B of Table 4.2 we observe that in no-reward groups, neither group leader bids nor group leader endorsements significantly influence the interest the borrower has to pay, except for slightly lower interest rates for credit grades D and HR. The effect is more pronounced for reward groups. The analysis by credit grade reveals that loans with a group leader bid or a group leader endorsement are associated with significantly smaller interest rates, in particular for the riskier credit grades. For example, borrowers with a loan in the credit grade HR pay on average 26.1% if the listing has neither a group leader bid nor a group leader endorsement, but only 24.2% if the group leader bids on the listing and only 24.3% if the group leader writes an endorsement.

From panel C of Table 4.2 we see that in no-reward groups, loans of the riskier credit grades E and HR have lower failure rates if they have a group leader bid or a group leader endorsement. By sharp contrast, loans in reward groups with a group leader bid or a group leader endorsement in general have significantly higher failure rates than loans without any of these two (18.9 / 19.0 vs. 15.7). This is the case for almost all credit grades. Apparently, group leader bids and group leader endorsements do not work as credible signals in reward groups.

Taken together, in both group types the success rates of listings with group leader bids and endorsements are much higher than for listings without group leader bids and endorsements. Yet, while in no-reward groups these two promotion mechanisms are associated with listings of good quality despite their bad credit grade E or HR, in reward groups failure rates are systematically increased for listings with a group leader bid or a group leader endorsement. Group leader bids and endorsements thus lead to adverse outcomes in reward groups. If this is due to adverse incentives for group leaders, then we should expect to see a change in behavior with a change in reward structure.

Group Leader Behavior Before and After the Elimination of Group Leader Rewards

We thus analyze next whether and how the change in reward structure affects the group leader behavior. Panel A of Figure 4.1 shows the weekly share of listings with at least one group leader bid in no-reward groups and in reward groups over the sample period. In no-reward groups, the share of listings with at least one group leader bid does not show any remarkable trend over the sample period. By sharp contrast, in reward groups this share decreases dramatically from about 40% to less than 10% once group leader rewards are eliminated.

Panel B of Figure 4.1 draws a similar picture for the other important mechanism: group leader endorsements. In particular, the share of listings with a group leader endorsement decreases significantly in reward groups from about 20% to less than 10% after the elimination of group leader rewards. The slight and rather slow increase of the respective share in the no-reward groups can be explained by the fact that friend endorsements were introduced only shortly before the beginning of our sample period (also see Figure 4.3), so that if nothing had changed – i.e. if group leader rewards had not been eliminated – we would have expected the same trend for no-reward groups and reward groups.

Table 4.3 confirms the results from Figure 4.1 by considering different credit grades. The results in panel A suggest that the share of listings with a group leader bid in no-reward groups does not change significantly after the elimination of group leader rewards for any credit grade. It remains at a level of about 45%. In strict contrast, the decrease in reward groups is significant for all credit grades, and it is most distinct for riskier credit grades. For example, it decreases from 34.7% to 3.9% for credit grade HR.

Panel B shows the respective results for the group leader endorsements. In no-reward groups, the share of listings with group leader endorsements increases on average after the elimination of group leader rewards, consistent with Figure 4.1. In contrast, in reward groups, the share of listings with a group leader endorsement decreases after the elimination of group leader rewards from 13.9% to 6.8%, which is especially due to the significant decrease in the corresponding shares of the high-risk listings with credit grades C, D, E and HR.

In sum, these results indicate that group leaders of reward groups significantly lower the effort they put into listings and in particular risky listings after the elimination of group leader rewards – as opposed to group leaders of no-reward groups who do not change their behavior. The resulting open question is how this change in behavior affects outcomes.

Figure 4.1: Weekly Share of Listings with a Group Leader Bid (Panel A) / with a Group Leader Endorsement (Panel B)

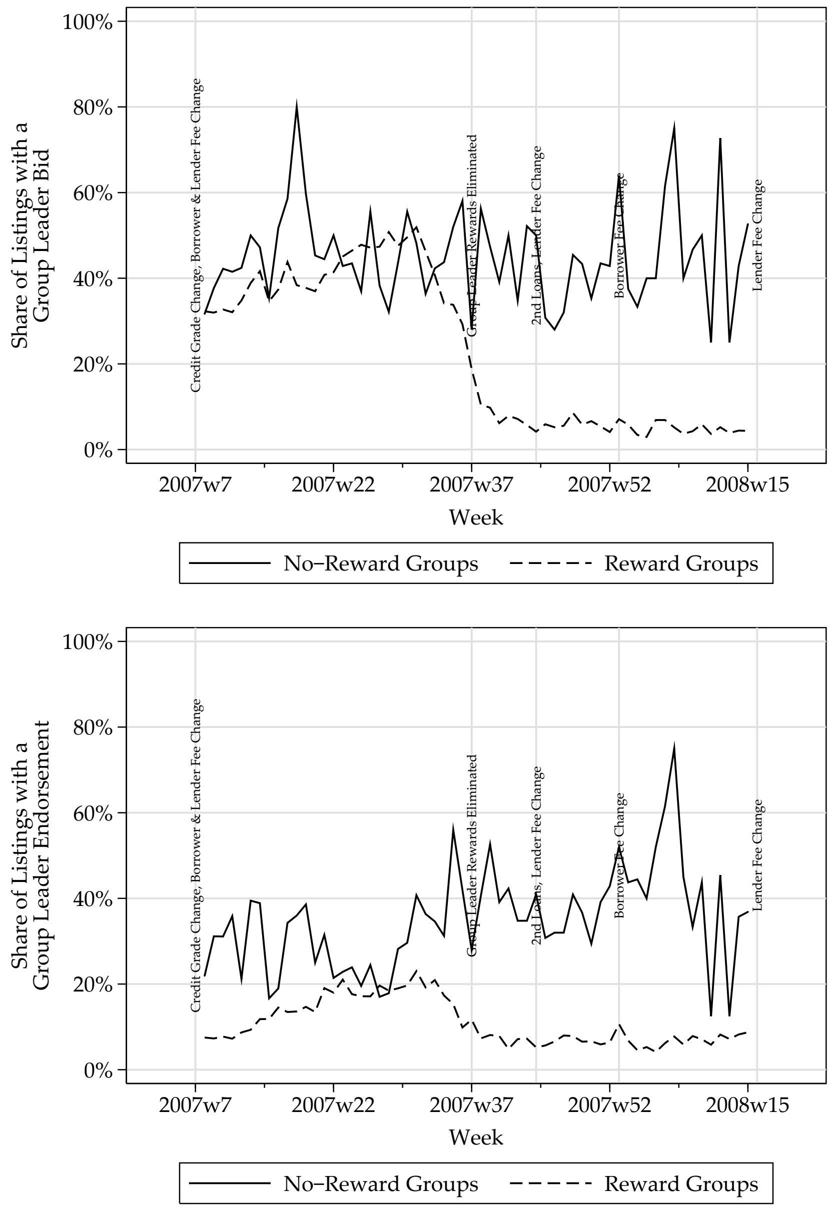


Table 4.3: Use of Group Leader Bids and Group Leader Endorsements

Panel A: Share of Listings with a Group Leader Bid

	No-Reward Groups (N=1,892)			Reward Groups (N=32,966)		
Credit Grade	**Before**	**After**	**t-statistic**	**Before**	**After**	**t-statistic**
AA/A	42.6%	42.4%	(0.02)	43.3%	24.0%	(6.70)***
B	44.2%	45.0%	(-0.10)	45.4%	15.1%	(12.37)***
C	52.2%	42.5%	(1.63)	42.7%	10.4%	(21.67)***
D	57.3%	52.0%	(0.90)	44.2%	5.9%	(37.25)***
E	45.0%	39.5%	(0.83)	37.6%	5.2%	(35.45)***
HR	40.1%	44.0%	(-0.84)	34.7%	3.9%	(54.03)***
Total	**46.5%**	**44.3%**	**(0.92)**	**38.8%**	**6.4%**	**(77.10)*** **

Panel B: Share of Listings with a Group Leader Endorsement

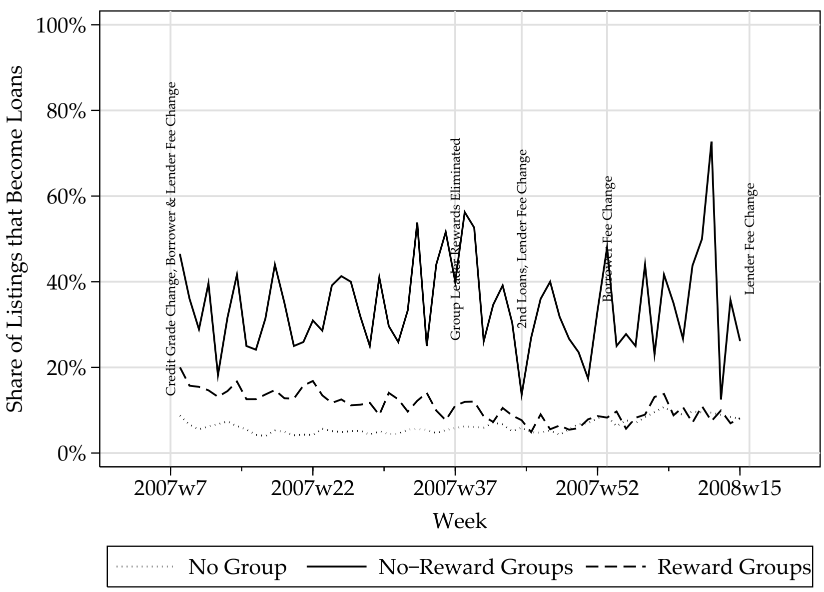
	No-Reward Groups (N=1,892)			Reward Groups (N=32,966)		
Credit Grade	**Before**	**After**	**t-statistic**	**Before**	**After**	**t-statistic**
AA/A	40.6%	34.3%	(1.06)	22.0%	19.8%	(0.85)
B	26.7%	40.0%	(-1.66)	20.1%	16.4%	(1.53)
C	27.8%	34.5%	(-1.20)	17.0%	9.9%	(5.26)***
D	30.5%	47.0%	(-2.84)***	16.4%	6.2%	(11.25)***
E	23.9%	38.3%	(-2.28)**	12.2%	6.3%	(7.46)***
HR	25.9%	44.6%	(-4.18)***	10.8%	4.6%	(12.93)***
Total	**29.1%**	**40.2%**	**(-4.72)*** **	**13.9%**	**6.8%**	**(18.97)*** **

In this table we report the share of listings (i.e. of requests for borrowing money) with at least one group leader bid (panel A) and the share of listings with a group leader endorsement (panel B) by group type and credit grade. T-statistics of the test on equality (before vs. after the elimination of group leader rewards) are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Effect of Change in Group Leader Behavior

A first price of evidence for the effect of the change in group leader behavior on outcomes is provided by Figure 4.2, which shows success rates of listings posted outside groups as well as of listings posted in no-reward groups and in reward groups. As shown before, success rates of listings in no-reward groups are generally the highest ones: they are significantly higher than those of listings in reward groups and those of listings posted outside groups. Success rates of listings in reward groups are also higher than those not posted in groups, but, most importantly for the purpose of this study, only before group leader rewards are eliminated and in a short transition period after the change.

Figure 4.2: Weekly Share of Successful Listings (Loans)



The changes in outcome patterns are analyzed in more detail in Table 4.4. Panel A of Table 4.4 shows that the overall success rate remains constant at 34.6% in no-reward groups before and after the elimination of group leader rewards. The results are also very similar for each of the different credit grades, with the exception of HR. In strict contrast to no-reward groups, success rates in reward groups decrease significantly from 13.4% to 8.6%. This decrease is particularly pronounced in the risky credit grades C to HR, while there is no significant change for the credit grades AA/A and B. This means that worse credit grades have a substantially lower chance of getting funded after the elimination of group leader rewards.

Table 4.4: Listing Success, Interest Rates, and Loan Performance Before and After Elimination of Group Leader Rewards

Panel A: Listing Success

	No-Reward Groups (N=1,892)			Reward Groups (N=32,966)		
Credit Grade	**Before**	**After**	**t-statistic**	**Before**	**After**	**t-statistic**
AA/A	59.9%	60.6%	(-0.12)	40.0%	41.0%	(-0.31)
B	47.7%	53.3%	(-0.67)	29.4%	29.1%	(0.10)
C	40.6%	40.7%	(-0.03)	24.1%	18.0%	(3.64)***
D	36.6%	40.0%	(-0.59)	17.7%	11.0%	(6.16)***
E	23.3%	25.9%	(-0.44)	9.2%	4.2%	(7.37)***
HR	19.3%	9.0%	(3.39)***	5.0%	3.0%	(5.31)***
Total	**34.6%**	**34.6%**	**(0.00)**	**13.4%**	**8.6%**	**(12.06)*** **

Panel B: Interest Rates

	No-Reward Groups (N=654)			Reward Groups (N=4,084)		
Credit Grade	**Before**	**After**	**t-statistic**	**Before**	**After**	**t-statistic**
AA/A	9.1%	9.7%	(-1.16)	11.3%	11.2%	(0.36)
B	12.5%	13.6%	(-1.70)*	14.9%	15.3%	(-0.90)
C	15.1%	16.3%	(-1.30)	17.4%	18.1%	(-1.52)
D	17.4%	18.4%	(-1.24)	20.1%	20.1%	(0.17)
E	21.3%	20.0%	(0.91)	23.9%	25.4%	(-1.79)*
HR	20.2%	21.7%	(-0.72)	24.5%	26.8%	(-3.07)***
Total	**15.1%**	**15.2%**	**(-0.09)**	**18.7%**	**18.9%**	**(-0.78)**

Panel C: Loan Performance

	No-Reward Groups (N=654)			Reward Groups (N=4,084)		
Credit Grade	**Before**	**After**	**t-statistic**	**Before**	**After**	**t-statistic**
AA/A	3.5	6.7	(6.37)***	9.0	8.0	(-2.56)**
B	7.3	7.3	(-0.06)	14.9	13.8	(-2.22)**
C	9.6	9.8	(0.25)	17.3	13.4	(-9.08)***
D	10.2	10.1	(-0.11)	17.9	11.2	(-17.32)***
E	14.2	13.7	(-0.42)	22.2	17.1	(-7.75)***
HR	24.3	14.2	(-7.80)***	26.2	22.5	(-6.12)***
Total	**11.6**	**9.5**	**(-6.61)*** **	**18.1**	**14.0**	**(20.43)*** **

In this table we report univariate results by group type and credit grade. We also distinguish whether the listing (i.e. the request for borrowing money) or the loan (i.e. the successfully and completely funded request for borrowing money) was created before or after the elimination of group leader rewards. Panel A shows success rates of listings by the different credit grades from AA/A (best) to HR (worst). Panel B shows the corresponding interest rates of loans. Panel C shows failure rates of loans (per 1,000 loan-days). In this panel, any payment which is not made on time is considered as a failure, so that failure events are late payments, charge-offs and defaults. T-statistics of the test on equality (before vs. after the elimination of group leader rewards) are reported in parentheses for both No-Reward Groups and Reward Groups. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Panel B of Table 4.4 suggests that interest rates do not significantly change after the elimination of group leader rewards, neither in no-reward groups nor in reward groups. The only exceptions are interest rates for credit grade B in no-reward groups and credit grades E and HR in reward groups, which pay slightly more after the change.

As shown in panel C of Table 4.4, failure rates in reward groups consistently decrease after the elimination of group leader rewards across all credit grades. The average decrease in failure rates of loans per 1,000 loan-days amounts to about 4. In the extreme case, failure rates decrease from 17.9 to 11.2 for credit grade D. In no-reward groups, no systematic pattern can be found. While failure rates increase for credit grades AA/A, they decrease for credit grade HR.

Taken together, these results show that no-reward groups work the same way before and after the elimination of group leader rewards. In contrast, reward groups work much better after the elimination of group leader rewards than before, as failure rates are substantially lower. A decrease in listing success along with a decrease in failure after the elimination of group leader rewards suggests that group leaders now much more carefully screen and choose the listings that are funded. An open question is why – before the elimination of group leader rewards – the listing success in reward groups is high despite the fact that the resulting loans also have a high likelihood of defaulting. This suggests that co-lenders do not fully foresee the consequences of the adverse incentives created by upfront rewards, most likely because of the short period between the creation of the webpage and the point of time when these lenders have to make their decisions.[6]

4.3.2 Multivariate Analysis

In order to determine the driving factors behind the results described above and to control for the joint influences, we now turn to the multivariate analysis.

Listing Success

Table 4.5 shows odds ratios of logistic regressions of listing success. In specification (1), we consider all listings, i.e. those posted in groups as well as those posted outside groups. Almost all covariates are highly significant and go into the expected direction: Listing success is decreasing in credit grade risk, debt-to-income ratio, and the number of historical and current records in the credit report; it is increasing in homeownership

[6]Lenders do not possess the full information that is used in this paper, as their decisions are made within the sample period, while the data for this paper cover the whole sample period.

and in income. Self-employed and in particular retired or unemployed borrowers face a particularly low funding probability. In terms of the listing characteristics, listing success is decreasing in the amount requested and increasing in the duration of the listing. Potential borrowers who decide to close their listing as soon as it is funded also exhibit higher chances to have their listing funded; obviously potential lenders tend to jump on these listings as there is a good chance to earn high interest rates given that one cannot be outbid.

Specification (1) considers all listings – independently of whether they are posted inside or outside groups – and shows that listings that are not posted in a group (*No Group*) or that are posted in a reward group (*Reward Group*) have significantly lower funding probabilities than those posted in no-reward groups, which is the reference group in all our regressions. Moreover, after the elimination of group leader rewards (*After*), listing success decreases.

In specifications (2) to (4) of Table 4.5, we concentrate on those listings that are posted in groups and analyze in particular the different group-specific variables.[7] The probability that the listing is funded increases significantly if the group leader requires the listing to be reviewed before it is posted in the group (*Listing Review Requirement*) or if the group leader offers help in designing the listing (*Group Leader Offers Help*). Vetting, i.e. the verification of the information by the group leader, seems surprisingly unimportant for the success of the listing. However, by far the most important group variables in terms of listing success are group leader bids and group leader endorsements at the top of specifications (2) to (4), which we analyze now more closely.

In specification (2), we include dummy variables for group leader bids and group leader endorsements into the regression and distinguish between *Only GL Bid*, *Only GL Endorsement* and *GL Bid & GL Endorsement*. Listings that have *GL Bid & GL Endorsement* exhibit particularly high funding probabilities. Listings with just one of these two elements are still about two to three times more likely to be funded than listings without any of these two. When comparing the coefficients for *Only GL Endorsement* and *Only GL Bid*, it may seem surprising at first sight that *Only GL Endorsement* – where there is no monetary commitment by the group leader at stake, i.e. where group leaders do not have "skin in the game" – has an even slightly higher positive influence on the funding probability than *Only GL Bid* has. We analyze this observation more carefully in the next specification.

In specification (3), we break down the influence of group leader bids and group leader endorsements for reward and no-reward groups. The results show that *Only GL*

[7] The results obtained with respect to the other covariates are robust across the different specifications.

Table 4.5: Listing Success – Multivariate Analysis

	All Listings		Only Listings in Groups					
	(1)		(2)		(3)		(4)	
GL Bids & GL Endorsements								
Only GL Bid			1.829***	(12.64)				
Only GL Bid: No-Reward					2.192***	(4.85)	2.172***	(4.80)
Only GL Bid: Reward					1.796***	(11.82)	1.772***	(11.53)
Only GL Endorsement			2.919***	(12.06)				
Only GL Endorsement: No-Reward					1.913**	(2.56)	1.916**	(2.56)
Only GL Endorsement: Reward					3.149***	(12.22)	3.157***	(12.24)
GL Bid & GL Endorsement			7.739***	(38.53)				
GL Bid & GL Endorsement: No-Reward					11.584***	(16.11)	11.580***	(16.11)
GL Bid & GL Endorsement: Reward					7.368***	(35.86)		
GL Bid & GL Endorsement: Reward, Before							7.038***	(33.89)
GL Bid & GL Endorsement: Reward, After							11.801***	(15.27)
Group Characteristics								
No Group	0.162***	(-29.83)						
Reward Group	0.414***	(-14.18)	0.573***	(-8.56)	0.669***	(-3.76)	0.661***	(-3.87)
Vetting			1.085	(1.40)	1.099	(1.61)	1.071	(1.15)
Listing Review Requirement			1.492***	(9.64)	1.494***	(9.65)	1.491***	(9.61)
Group Leader Offers Help			1.375***	(5.08)	1.336***	(4.56)	1.334***	(4.53)
Listing Characteristics								
After	0.857***	(-6.09)	0.790***	(-4.50)	0.781***	(-4.70)	0.740***	(-5.41)
Amount Requested (in $1,000)	0.887***	(-57.39)	0.894***	(-29.83)	0.893***	(-29.82)	0.893***	(-29.90)
Duration	1.063***	(11.67)	1.036***	(3.70)	1.038***	(3.82)	1.038***	(3.82)
Listing Closed As Soon As Funded	1.140***	(5.13)	0.939	(-1.38)	0.938	(-1.40)	0.938	(-1.40)
Borrower Characteristics								
Credit Grade: B	0.612***	(-12.81)	0.663***	(-5.20)	0.656***	(-5.33)	0.658***	(-5.29)
Credit Grade: C	0.302***	(-32.71)	0.426***	(-11.91)	0.419***	(-12.10)	0.422***	(-11.99)
Credit Grade: D	0.153***	(-47.83)	0.237***	(-19.44)	0.234***	(-19.61)	0.236***	(-19.48)
Credit Grade: E	0.060***	(-56.96)	0.102***	(-26.60)	0.100***	(-26.73)	0.101***	(-26.61)
Credit Grade: HR	0.027***	(-71.02)	0.055***	(-33.19)	0.055***	(-33.29)	0.055***	(-33.21)
Debt-to-Income Ratio	0.900***	(-9.89)	0.967**	(-2.48)	0.967**	(-2.52)	0.966***	(-2.60)
Is Borrower Home Owner	1.167***	(6.22)	1.160***	(3.45)	1.163***	(3.52)	1.164***	(3.53)
$1-24,999	1.316***	(2.70)	0.827	(-1.20)	0.830	(-1.17)	0.830	(-1.18)
$25,000-49,999	1.895***	(6.35)	1.233	(1.32)	1.231	(1.31)	1.234	(1.33)
$50,000-74,999	2.391***	(8.54)	1.658***	(3.14)	1.657***	(3.14)	1.661***	(3.15)
$75,000-99,999	3.000***	(10.42)	2.038***	(4.23)	2.040***	(4.23)	2.049***	(4.26)
$100,000	3.409***	(11.42)	2.432***	(5.12)	2.434***	(5.12)	2.451***	(5.16)
Part-Time	1.000	(0.00)	0.864	(-1.40)	0.854	(-1.50)	0.853	(-1.51)
Self-Employed	0.924*	(-1.86)	1.074	(1.00)	1.070	(0.94)	1.071	(0.96)
Retired	0.643***	(-5.72)	0.692***	(-2.84)	0.686***	(-2.90)	0.688***	(-2.88)
Not Employed	0.632***	(-3.18)	0.597**	(-2.38)	0.591**	(-2.43)	0.593**	(-2.41)
Current Delinquencies	0.917***	(-14.53)	0.961***	(-4.91)	0.961***	(-4.91)	0.962***	(-4.86)
Delinquencies Last 7 Years	0.995***	(-5.07)	0.997	(-1.63)	0.997	(-1.62)	0.997	(-1.59)
Public Records Last 10 Years	0.970**	(-2.38)	0.959**	(-1.97)	0.959**	(-1.97)	0.958**	(-2.00)
Total Credit Lines	0.993***	(-5.57)	0.994***	(-3.22)	0.993***	(-3.30)	0.993***	(-3.31)
Inquiries Last 6 Months	0.974***	(-8.93)	0.986***	(-3.29)	0.986***	(-3.24)	0.986***	(-3.19)
Amount Delinquent (in $1,000)	0.993***	(-2.89)	0.991**	(-2.46)	0.991**	(-2.46)	0.990**	(-2.51)
Public Records Last 12 Months	1.084*	(1.88)	1.087	(1.21)	1.089	(1.24)	1.091	(1.27)
Current Credit Lines	1.004	(0.59)	1.034***	(3.34)	1.033***	(3.31)	1.033***	(3.29)
Open Credit Lines	0.973***	(-4.25)	0.957***	(-4.09)	0.957***	(-4.04)	0.958***	(-4.02)
Revolving Credit Balance (in $1,000)	1.000	(1.09)	0.999	(-1.31)	0.999	(-1.40)	0.999	(-1.40)
Bankcard Utilization	1.081**	(2.43)	1.005	(0.09)	1.003	(0.06)	1.005	(0.10)
Months in Current Occupation	1.000***	(-2.62)	0.999**	(-2.34)	0.999**	(-2.28)	0.999**	(-2.31)
N	153,541		34,858		34,858		34,858	
pseudo R^2	0.258		0.275		0.276		0.276	

In this table we report odds ratios of the logistic regression of funding success, i.e. the exponentiated regression coefficients. Coefficients larger (respectively smaller) than 1 indicate relatively higher (respectively smaller) success probabilities than in the reference group. In specification (1) all listings (i.e. all requests for borrowing money) are considered, in specifications (2) to (4) only group listings are analyzed. Specification (2) reports the overall effect of a group leader bid and / or a group leader endorsement on listing success. Specification (3) additionally distinguishes whether the group leader bid and / or the group leader endorsement occurs in a listing in a no-reward group or in a reward group. Specification (4) compares the joint effect of a group leader bid and a group leader endorsement before and after the elimination of group leader rewards on listing success in the reward groups. Note that in this specification, the difference between the regression coefficients of "GL Bid & GL Endorsement: Reward, Before" and "GL Bid & GL Endorsement: Reward, After" is significant at 1%. The reference is AA/A-listings before the elimination of group leader rewards in no-reward groups without a group leader bid or a group leader endorsement. T-statistics are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Bid, *Only GL Endorsement* and *GL Bid & GL Endorsement* work in the same way in reward and no-reward groups. However, *Only GL Endorsement* works particularly well in reward groups, while *Only GL Bid* works better in no-reward groups. The larger coefficient for *Only GL Endorsement* in specification (2) is thus solely due to its higher listing success in reward groups. We will later analyze whether these endorsements eventually also lead to loans with lower failure rates, or whether the group leader simply persuades potential lenders to participate in a loan so that he can earn the upfront reward associated with a successful listing.

Finally, specification (4) constitutes the key part of our analysis and employs a difference-in-difference methodology with two sources of identifying variation: (i) the time before and after the removal of group leader rewards, (ii) the distinction between listings inside and outside reward groups. Our inference is based on evaluating whether reward groups perform differently after the elimination of group leader rewards. It shows that after this event the influence of the combination of a group leader bid and a group leader endorsement in the reward groups is significantly higher than before.[8] The result indicates that – after the elimination of group leader rewards – potential lenders trust much more than before the correctness of the group leader's signal that comes from his bid and endorsement. This suggests that after this change, lenders might be less concerned about the group leader behaving opportunistically and promoting listings only for his own benefit.

Interest Rates of Loans

In order to determine the influence of the different variables on the interest rates that borrowers have to pay to the lenders if their listing is funded, we run Tobit regressions of this interest rate (in percent) on the same independent variables as in the regressions in Table 4.5. Table 4.6 reports the results, where the dependent variable is truncated at left at 0% and at right at 35%, which is the maximum interest rate possible on Prosper.com.[9] Naturally, the sample is restricted to those listings that are completely funded and therefore become loans.

The interest rate of loans in the reference group, which are AA/A-loans, is about 5%. As before, most covariates are significant and have the expected signs. The borrower's credit grade is by far the most important influencing factor for the interest rate charged

[8] Due to the high correlation of group leader bids and group leader endorsements and the resulting low sample size for *Only GL Bid* and *Only GL Endorsement* after the elimination of group leader rewards, we do not distinguish the two variables *Only GL Bid* and *Only GL Endorsement* in the reward groups between before and after the elimination of group leader rewards.

[9] OLS regression results differ only marginally and are therefore not reported here.

Table 4.6: Interest Rates – Multivariate Analysis

	All Loans		Only Loans in Groups					
	(1)		(2)		(3)		(4)	
GL Bids & GL Endorsements								
Only GL Bid			-0.713***	(-5.07)				
Only GL Bid: No-Reward					-1.320***	(-3.22)	-1.285***	(-3.14)
Only GL Bid: Reward					-0.642***	(-4.35)	-0.595***	(-4.01)
Only GL Endorsement			0.213	(0.95)				
Only GL Endorsement: No-Reward					-0.067	(-0.12)	-0.052	(-0.09)
Only GL Endorsement: Reward					0.242	(0.99)	0.261	(1.07)
GL Bid & GL Endorsement			-0.886***	(-6.27)				
GL Bid & GL Endorsement: No-Reward					-1.076***	(-3.11)	-1.061***	(-3.06)
GL Bid & GL Endorsement: Reward					-0.878***	(-5.90)		
GL Bid & GL Endorsement: Reward, Before							-0.755***	(-4.90)
GL Bid & GL Endorsement: Reward, After							-1.807***	(-5.24)
Group Characteristics								
No Group	2.060***	(12.76)						
Reward Group	1.342***	(8.14)	1.263***	(8.41)	1.010***	(3.45)	1.052***	(3.60)
Vetting			-0.501***	(-3.50)	-0.496***	(-3.44)	-0.421***	(-2.88)
Listing Review Requirement			0.118	(0.98)	0.128	(1.07)	0.130	(1.09)
Group Leader Offers Help			-0.721***	(-4.72)	-0.712***	(-4.62)	-0.713***	(-4.63)
Listing Characteristics								
After	1.345***	(15.42)	1.499***	(10.59)	1.500***	(10.58)	1.691***	(10.88)
Amount Requested (in $1,000)	0.253***	(36.49)	0.290***	(29.02)	0.290***	(29.03)	0.291***	(29.18)
Duration	-0.007	(-0.39)	0.009	(0.37)	0.008	(0.32)	0.009	(0.36)
Listing Closed As Soon As Funded	3.286***	(37.07)	2.961***	(22.85)	2.971***	(22.90)	2.977***	(22.96)
Borrower Characteristics								
Credit Grade: B	3.619***	(31.20)	2.896***	(15.69)	2.895***	(15.67)	2.880***	(15.60)
Credit Grade: C	6.299***	(54.49)	5.732***	(33.47)	5.729***	(33.36)	5.706***	(33.23)
Credit Grade: D	9.586***	(74.34)	8.634***	(47.41)	8.635***	(47.36)	8.611***	(47.23)
Credit Grade: E	13.580***	(80.37)	12.249***	(54.57)	12.241***	(54.51)	12.206***	(54.33)
Credit Grade: HR	13.420***	(75.66)	12.917***	(55.77)	12.916***	(55.76)	12.892***	(55.67)
Debt-to-Income Ratio	0.157***	(4.70)	0.162***	(4.45)	0.161***	(4.43)	0.166***	(4.57)
Is Borrower Home Owner	-0.152*	(-1.82)	-0.500***	(-4.35)	-0.499***	(-4.34)	-0.502***	(-4.37)
$1-24,999	0.220	(0.64)	0.971**	(2.27)	0.966**	(2.26)	0.956**	(2.24)
$25,000-49,999	-0.340	(-1.00)	0.455	(1.08)	0.456	(1.08)	0.449	(1.06)
$50,000-74,999	-0.473	(-1.38)	0.232	(0.54)	0.235	(0.55)	0.226	(0.53)
$75,000-99,999	-0.733**	(-2.08)	-0.180	(-0.41)	-0.181	(-0.41)	-0.194	(-0.44)
$100,000	-1.132***	(-3.16)	-0.579	(-1.27)	-0.580	(-1.27)	-0.594	(-1.31)
Part-Time	-0.423**	(-2.19)	-0.034	(-0.12)	-0.041	(-0.15)	-0.047	(-0.17)
Self-Employed	0.221	(1.55)	0.145	(0.75)	0.136	(0.71)	0.132	(0.69)
Retired	0.129	(0.49)	-0.258	(-0.72)	-0.246	(-0.68)	-0.248	(-0.69)
Not Employed	0.605	(1.18)	1.125*	(1.81)	1.123*	(1.81)	1.095*	(1.76)
Current Delinquencies	0.072***	(4.15)	0.069***	(3.28)	0.068***	(3.28)	0.068***	(3.28)
Delinquencies Last 7 Years	0.025***	(7.07)	0.021***	(4.57)	0.020***	(4.54)	0.020***	(4.55)
Public Records Last 10 Years	0.203***	(4.70)	0.224***	(3.70)	0.224***	(3.70)	0.226***	(3.74)
Total Credit Lines	0.019***	(4.83)	0.013**	(2.48)	0.014**	(2.53)	0.014**	(2.53)
Inquiries Last 6 Months	0.141***	(14.18)	0.076***	(6.16)	0.076***	(6.14)	0.075***	(6.09)
Amount Delinquent (in $1,000)	0.018***	(3.14)	0.015	(1.55)	0.015	(1.58)	0.016	(1.63)
Public Records Last 12 Months	0.445***	(2.83)	0.179	(0.83)	0.177	(0.82)	0.171	(0.79)
Current Credit Lines	-0.054***	(-2.59)	-0.028	(-1.02)	-0.029	(-1.06)	-0.028	(-1.02)
Open Credit Lines	0.054**	(2.40)	0.023	(0.80)	0.024	(0.81)	0.022	(0.76)
Revolving Credit Balance (in $1,000)	0.001	(1.29)	0.004**	(2.00)	0.004**	(1.98)	0.004*	(1.95)
Bankcard Utilization	0.416***	(3.73)	0.449***	(3.09)	0.445***	(3.07)	0.443***	(3.05)
Months in Current Occupation	0.001	(0.97)	0.001	(0.91)	0.001	(0.93)	0.001	(0.98)
Constant	5.087***	(12.68)	5.817***	(11.73)	6.053***	(10.99)	5.957***	(10.81)
N	12,183		4,738		4,738		4,738	
pseudo R^2	0.160		0.180		0.180		0.180	

In this table we report the regression coefficients from Tobit regressions of the lender interest rate of loans (i.e. of successfully and completely funded requests for borrowing money). In specification (1) all loans are considered, in specifications (2) to (4) only group loans are analyzed. Specification (2) reports the overall effect of a group leader bid and / or a group leader endorsement on the borrower interest rate. Specification (3) additionally distinguishes whether the group leader bid and / or the group leader endorsement occurs in a loan in a no-reward group or in a reward group. Specification (4) compares the joint effect of a group leader bid and a group leader endorsement before and after the elimination of group leader rewards on the borrower interest rate of loans in the reward groups. Note that in this specification, the difference between the regression coefficients of "GL Bid & GL Endorsement: Reward, Before" and "GL Bid & GL Endorsement: Reward, After" is significant at 1%. The reference is AA/A-loans before the elimination of group leader rewards in no-reward groups without a group leader bid or a group leader endorsement. T-statistics are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

to the borrower. Apart from that, the borrower interest rate is increasing in the debt-to-income ratio and in the number of historical and current records in the credit report. It is also decreasing in income, although this effect becomes insignificant if only group loans in specifications (2) to (4) are considered. Furthermore, a higher amount requested typically increases the interest rate. The interest rate increases by about 3% if the borrower chooses that the listing shall be closed as soon as it is completely funded; the interest rates cannot be bid down in this case.

Specification (1) shows that interest rates of loans funded outside groups (*No Group*) or in reward groups (*Reward Group*) are higher than those of loans in no-reward groups. Specification (2) shows that loans originated from listings with *Only GL Bid* benefit from particularly low interest rates, and interest rates are even lower for loans with *GL Bid & GL Endorsement*. We also find that the interest rate of the loan is significantly lower if the group leader claims to verify additional information from the borrower (*Vetting*) or if the group leader offers help in designing the listing (*Group Leader Offers Help*).

Specification (3) shows the results for reward and no-reward groups. Loans with *Only GL Endorsement* do not benefit from significantly lower interest rates. Otherwise, group leader bids and endorsements lead to lower interest rates both in reward and no-reward groups.

Finally, from specification (4), which uses again a difference-in-difference methodology, we deduce that after the elimination of group leader rewards, the interest rate of loans with *GL Bid & GL Endorsement* in reward groups is about 1% smaller than before. This result indicates that after this event, group leader bids and group leader endorsements have a significantly higher influence on the resulting interest rate in this group type. This suggests again that the signal of a group leader bid and endorsement is much more credible after the elimination of group leader rewards than before.

Loan Performance

In order to analyze the determinants of loan performance, we specify Cox proportional hazards models with the same independent variables as before. The underlying assumption of the models is that the coefficients are not time-varying, i.e. the importance of a variable for the probability of defaulting or being late is constant over time.[10] Loans are exposed to the process from the time they are originated until they are either completely paid back, they default or their data runs out. The results of the Cox proportional haz-

[10] If e.g. a loan with credit grade HR is more susceptible to have a failure than a loan of the reference group AA/A, the strength of this relationship does not depend on time. Thus, for example, the HR-loan does not become more susceptible to fail over time, compared to the AA/A-loan.

ards models are reported in Table 4.7.

Specification (1) of Table 4.7 shows that hazard rates are increasing in the credit grade risk and the debt-to-income ratio. Borrowers who use their bankcard exhibit lower hazard rates. Hazard rates are decreasing in income, whereas borrowers who are unemployed or retired have higher hazard rates. In terms of the listing characteristics, hazard rates are increasing in the loan amount. Furthermore, if the listing has a short duration or if it is closed as soon as it is funded, the corresponding loan is potentially exposed to a higher hazard rate. Together, this suggests that borrowers in urgent need of money exhibit higher hazard rates. For the key variables of interest, the group type significantly influences hazard rates even after controlling for other factors. Loans in reward groups (*Reward Group*) and loans resulting from listings posted outside groups (*No Group*) exhibit significantly higher hazard rates than loans in no-reward groups as the reference group. This result confirms the evidence from the univariate analysis.

The results in specifications (2) to (5) suggest that hazard rates are also reduced if the group leader verifies the information provided (*Vetting*) or if he generally offers help in designing the listing (*Group Leader Offers Help*). Most importantly for the purpose of this study, specification (2) shows that while *Only GL Bid* is insignificant in explaining the failure rate of a loan, the opposite is the case for *Only GL Endorsement* or the combination *GL Bid & GL Endorsement*, which increase failure rates. Obviously, group leader endorsements do not work properly as a signal of good listing quality.

From specification (3) we see that this is only a problem in reward groups, whereas in no-reward groups *Only GL Bid*, *Only GL Endorsement* as well as the combination *GL Bid & GL Endorsement* significantly lower the hazard rate of the loan. One may wonder whether before the elimination of group leader rewards it is profitable for the group leaders of reward groups to promote listings in their groups by placing a group leader bid on them. Further analysis shows that in this time period the group leader rewards more than compensate for the slightly higher failure rates in these groups.[11]

Most importantly, the influence of the elimination of group leader rewards on loan performance in reward groups can be deduced from the difference-in-difference specification (4): while before this policy change the combination of *GL Bid & GL Endorsement*

[11] To be specific, we calculate the median internal rate of return (IRR) of three different investments the group leader can make: (i) investment in a listing in her reward group by placing a group leader bid, (ii) investment in a listing in a no-reward group and (iii) investment in a listing not posted in any group. The median IRRs of investments (ii) and (iii) are negative with -22.4% and -37.0% as most loans are not yet paid back completely. Only the median IRR of investment (i) is already positive with 7.2% – due to the additional reward the group leader obtains. This clearly shows that it is profitable for the group leader of a reward group to promote listings in her group so that she obtains the group leader reward.

Table 4.7: Loan Performance – Multivariate Analysis

	All Loans		Only Loans in Groups							
	(1)		(2)		(3)		(4)		(5)	
GL Bids & GL Endorsements										
Only GL Bid			0.998	(-0.14)						
Only GL Bid: No-Reward					0.906*	(-1.85)	0.914*	(-1.68)	0.951	(-0.94)
Only GL Bid: Reward					1.001	(0.05)	1.013	(0.90)	1.014	(0.97)
Only GL Endorsement			1.106***	(4.25)						
Only GL Endorsement: No-Reward					0.814**	(-2.35)	0.816**	(-2.33)	0.847*	(-1.89)
Only GL Endorsement: Reward					1.124***	(4.79)	1.128***	(4.94)	1.134***	(5.16)
GL-B & GL-E			1.105***	(7.23)						
GL-B & GL-E: No-Reward					0.841***	(-3.80)	0.845***	(-3.71)		
GL-B & GL-E: No-Reward, $\leq$ 33%									0.950	(-1.12)
GL-B & GL-E: No-Reward, $>$ 33%									0.337***	(-8.73)
GL-B & GL-E: Reward					1.125***	(8.39)				
GL-B & GL-E: Reward, Before							1.154***	(9.94)		
GL-B & GL-E: Reward, Before, $\leq$ 33%									1.172***	(10.92)
GL-B & GL-E: Reward, Before, $>$ 33%									0.821***	(-3.85)
GL-B & GL-E: Reward, After							0.823***	(-4.60)		
GL-B & GL-E: Reward, After, $\leq$ 33%									0.869***	(-3.30)
GL-B & GL-E: Reward, After, $>$ 33%									0.084***	(-4.95)
Group Characteristics										
No Group	1.307***	(15.12)								
Reward Group	1.419***	(20.02)	1.425***	(19.34)	1.172***	(4.04)	1.182***	(4.26)	1.225***	(5.10)
Vetting			0.865***	(-9.61)	0.856***	(-10.23)	0.874***	(-8.84)	0.882***	(-8.23)
Listing Review Requirement			0.994	(-0.51)	0.997	(-0.24)	0.997	(-0.26)	0.993	(-0.64)
Group Leader Offers Help			0.947***	(-3.53)	0.957***	(-2.82)	0.957***	(-2.83)	0.941***	(-3.91)
Listing Characteristics										
After	0.836***	(-20.37)	0.825***	(-11.83)	0.830***	(-11.48)	0.883***	(-6.96)	0.884***	(-6.93)
Amount Requested (in $1,000)	1.062***	(89.84)	1.061***	(60.08)	1.061***	(60.10)	1.062***	(60.38)	1.061***	(59.57)
Duration	0.983***	(-10.81)	0.979***	(-8.73)	0.978***	(-8.97)	0.979***	(-8.87)	0.979***	(-8.82)
Listing Closed As Soon As Funded	1.357***	(40.29)	1.171***	(13.44)	1.172***	(13.47)	1.173***	(13.51)	1.174***	(13.62)
Borrower Characteristics										
Credit Grade: B	1.747***	(40.83)	1.774***	(24.59)	1.773***	(24.57)	1.770***	(24.51)	1.764***	(24.35)
Credit Grade: C	2.305***	(62.86)	2.330***	(39.21)	2.333***	(39.23)	2.318***	(38.92)	2.305***	(38.65)
Credit Grade: D	2.792***	(72.11)	2.627***	(43.00)	2.633***	(43.08)	2.621***	(42.87)	2.604***	(42.56)
Credit Grade: E	3.812***	(81.09)	3.757***	(52.53)	3.760***	(52.54)	3.729***	(52.20)	3.717***	(52.06)
Credit Grade: HR	4.741***	(92.24)	5.019***	(63.39)	5.030***	(63.46)	4.992***	(63.14)	4.977***	(62.99)
Debt-to-Income Ratio	1.017***	(6.43)	1.022***	(6.48)	1.022***	(6.64)	1.023***	(6.82)	1.024***	(7.13)
Is Borrower Home Owner	1.151***	(17.70)	1.109***	(9.32)	1.110***	(9.36)	1.111***	(9.48)	1.108***	(9.20)
$1-24,999	1.126***	(3.76)	1.122**	(2.50)	1.118**	(2.42)	1.117**	(2.39)	1.107**	(2.21)
$25,000-49,999	1.074**	(2.29)	1.050	(1.08)	1.051	(1.09)	1.050	(1.06)	1.036	(0.77)
$50,000-74,999	0.939**	(-2.01)	0.938	(-1.40)	0.937	(-1.40)	0.935	(-1.45)	0.928	(-1.62)
$75,000-99,999	0.935**	(-2.08)	0.986	(-0.31)	0.985	(-0.31)	0.984	(-0.33)	0.971	(-0.63)
$100,000	0.827***	(-5.74)	0.855***	(-3.23)	0.852***	(-3.29)	0.847***	(-3.40)	0.840***	(-3.58)
Part-Time	0.991	(-0.48)	1.122***	(4.04)	1.131***	(4.30)	1.132***	(4.35)	1.120***	(3.98)
Self-Employed	1.106***	(7.82)	0.952***	(-2.68)	0.952***	(-2.63)	0.951***	(-2.69)	0.948***	(-2.85)
Retired	1.119***	(4.71)	1.315***	(8.79)	1.315***	(8.78)	1.317***	(8.83)	1.324***	(9.00)
Not Employed	1.333***	(6.58)	1.326***	(4.54)	1.324***	(4.50)	1.319***	(4.44)	1.351***	(4.81)
Current Delinquencies	1.023***	(20.59)	1.025***	(17.85)	1.025***	(18.02)	1.025***	(17.98)	1.025***	(17.83)
Delinquencies Last 7 Years	0.998***	(-7.21)	0.998***	(-5.25)	0.998***	(-5.41)	0.998***	(-5.43)	0.998***	(-4.99)
Public Records Last 10 Years	1.046***	(14.87)	1.074***	(15.05)	1.075***	(15.14)	1.076***	(15.35)	1.076***	(15.28)
Total Credit Lines	1.006***	(16.79)	1.005***	(11.05)	1.005***	(11.02)	1.005***	(11.14)	1.005***	(11.19)
Inquiries Last 6 Months	1.047***	(71.17)	1.043***	(49.79)	1.043***	(49.76)	1.043***	(49.76)	1.043***	(49.91)
Amount Delinquent (in $1,000)	1.000	(-0.31)	1.003***	(4.43)	1.003***	(4.45)	1.003***	(4.56)	1.003***	(4.44)
Public Records Last 12 Months	0.962***	(-3.09)	0.947***	(-2.98)	0.948***	(-2.96)	0.945***	(-3.08)	0.946***	(-3.07)
Current Credit Lines	1.002	(1.17)	1.005**	(1.99)	1.006**	(2.18)	1.005**	(2.10)	1.005**	(1.98)
Open Credit Lines	0.986***	(-6.79)	0.987***	(-4.45)	0.987***	(-4.61)	0.987***	(-4.59)	0.988***	(-4.41)
Revolving Credit Balance (in $1,000)	1.000***	(5.24)	1.001***	(6.11)	1.001***	(6.16)	1.001***	(6.05)	1.001***	(6.11)
Bankcard Utilization	0.935***	(-6.83)	0.924***	(-5.95)	0.924***	(-5.92)	0.923***	(-5.99)	0.924***	(-5.93)
Months in Current Occupation	1.000***	(-2.80)	1.000	(-1.61)	1.000*	(-1.77)	1.000*	(-1.81)	1.000**	(-1.99)
N	374,235		161,000		161,000		161,000		161,000	

In this table we report the exponentiated regression coefficients obtained from a Cox Proportional Hazards Model. Any payment which is not made on time is considered as a failure, so that failure events are late payments, charge-offs and defaults. In specification (1) all loans (i.e. all successfully and completely funded requests for borrowing money) are considered, in specifications (2) to (5) only group loans are analyzed. Specification (2) reports the overall effect of a group leader bid and / or a group leader endorsement on the failure probability of loans. Specification (3) additionally distinguishes whether the group leader bid and / or the group leader endorsement occurs in a loan in a no-reward group or in a reward group. Specification (4) compares the joint effect of a group leader bid and a group leader endorsement before and after the elimination of group leader rewards on the failure probability of loans in the reward groups. Note that in this specification, the difference between the regression coefficients of "GL Bid & GL Endorsement: Reward, Before" and "GL Bid & GL Endorsement: Reward, After" is significant at 1%. Finally, specification (5) analyzes whether before the elimination of group leader rewards, the group leader participates with more than 33% of the loan amount in the loan, if she places a bid and an endorsement on the listing (i.e. whether she "has skin in the game"). The reference is AA/A-loans before the elimination of group leader rewards in no-reward groups without a group leader bid or a group leader endorsement. T-statistics are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

hints at a ceteris paribus higher hazard rate (coefficient of 1.154), after this event the hazard rate is significantly smaller not only than before the change but also than the benchmark of 1 (coefficient of 0.823). Consequently, the results suggest that – before the elimination of group leader rewards – group leaders of reward groups overpromote bad listings with the help of group leader bids and especially group leader endorsements, which lead to higher failure rates for these types of loans. In contrast, after this policy change, the mechanism works properly as the group leader has now no incentive any more to behave opportunistically.

The evidence so far suggests that rewards give group leaders an incentive to promote and bid even on bad listings as these rewards more than offset the losses due to the higher likelihood of failure. This behavior changes once the reward is eliminated, which changes the group leaders' trade-off between rewards and losses. An alternative way to align incentives, i.e. to make group leaders screen listings very carefully, is that – even before the elimination of group leader rewards – group leaders participate to a large fraction in the loan and thus have substantial skin in the game. We therefore further differentiate in specification (5) whether a group leader participates in more or less than 33% of the loan.[12] The results show that the failure rates decrease substantially when the group leader participates in more than 33% of the loan; this holds for no-reward groups as well as reward groups before and after the elimination of group leader rewards. However, only in reward groups before the event, the failure rate is higher than 1 if the group leader participates in less than 33% of the loan. This means that the potential losses in this case are not high enough to outweigh the rewards. Or, interpreted differently, only a large commitment and thus substantial skin in the game induces a group leader to carefully screen borrowers and promote the creditworthy listings, even

[12] The threshold of 33% is obtained as follows: A listing yields a negative payoff to a regular bidder under the following simplified condition: $-\alpha+\alpha I(1-p)+\alpha(1-p)<0$, where α = share of the loan amount supplied by this bidder, I = interest rate obtained, p = probability of default. The recovery rate is assumed to be zero. This can be simplified to $-\alpha(Ip+p-I)<0$, so that $\alpha>0$ implies $(Ip+p-I)>0$ for a listing with a negative payoff. Suppose the group leader knows p and I from historical data. To make it profitable for him to still bid on a listing with a negative payoff, group leader fees and upfront payment have to outweigh the loss: $F(1-p)+U>\alpha(Ip+p-I)$, where F = group leader fee (interest rate paid on the full loan amount), and U = upfront payment to the group leader (relative to the loan amount). Since $(Ip+p-I)>0$ as before, $(F(1-p)+U)/(Ip+p-I)>\alpha$ yields an upper bound for a profitable group leader bid on this listing. For each credit grade we compute the critical value α according to this last formula. As an example, consider a borrower with the credit grade B in a reward group. For this borrower, we have the average interest rate $I=15\%$, the probability of default $p=18\%$, the group leader fee $F=2\%$ and the upfront fee $U=0.5\%$. According to the formula above this yields a cutoff criterion of $(0.02x(1-0.18)+0.005)/(0.15x0.18+0.18-0.15)=0.37>\alpha$. Consequently, the group leader should not participate in more than 37% of B-loans in which a regular bidder would lose money. The resulting overall critical value of 33% is the weighted average over these critical values of the credit grades.

if he can earn rewards. The coefficient of 0.821 in this case is almost identical to that of 0.823 in specification (4), which captures the failure probability after the elimination of the rewards. These results suggest that a high bid by the group leader serves indeed as a signal about the quality of screening, as the other lenders correctly assume that a higher participation leads to more skin in the game and thus a more careful screening process.

4.4 Conclusion

Lenders and borrowers in markets without financial intermediaries with skin in the game are confronted with even more substantial information asymmetries than agents in traditional lending markets. An important open question is thus how these markets can work properly and efficiently, i.e. give borrowers access to credit at rates that incorporate their risk of default and protect them against unscrupulous lending. The analyses in this paper for the online social lending market show that information asymmetries can be alleviated by a system in which group leaders credibly signal borrower quality to other lenders by endorsing and submitting bids for carefully screened borrower listings.

The functioning of groups is severely impaired when group leaders are rewarded for successful listings. We thus show that online social lending platforms are indeed right in eliminating or not using rewards that give group leaders adverse incentives to promote non-sustainable loans. After the elimination of group leader rewards, reward groups work much better and provide the correct incentive structures to the group leader. Even before the elimination of rewards, these groups work well if the group leader puts his money where his mouth is and participates to a substantial fraction in a loan. Listings are then promoted only if the group leader trusts their quality. These results show that only a considerable fraction of the loan retained by group leaders in reward groups induces them to efficiently and responsibly screen loan listings. This result does not imply that group leaders should generally not be rewarded. Rather some other mechanisms might be considered, e.g. the group leader might obtain a small lump sum from the borrower once the loan is completely paid back.

In sum, this paper has at least two important implications. First, the results have direct relevance for the question of how to protect retail customers in the substantially growing online social lending markets. Second, while they cannot be simply generalized to other financial markets in which consumer protection is also of vital interest, they provide evidence from a very clean experiment and show that proper incentives are crucial for giving borrowers access to credit and to induce lenders to carefully screen loan applicants. It is an open question for future research how these findings can be more generally applied to consumer protection in traditional financial and lending industry as well as in more unconventional ways of access to finance such as micro lending.

4.A Appendix: Timeline and Variable Definitions

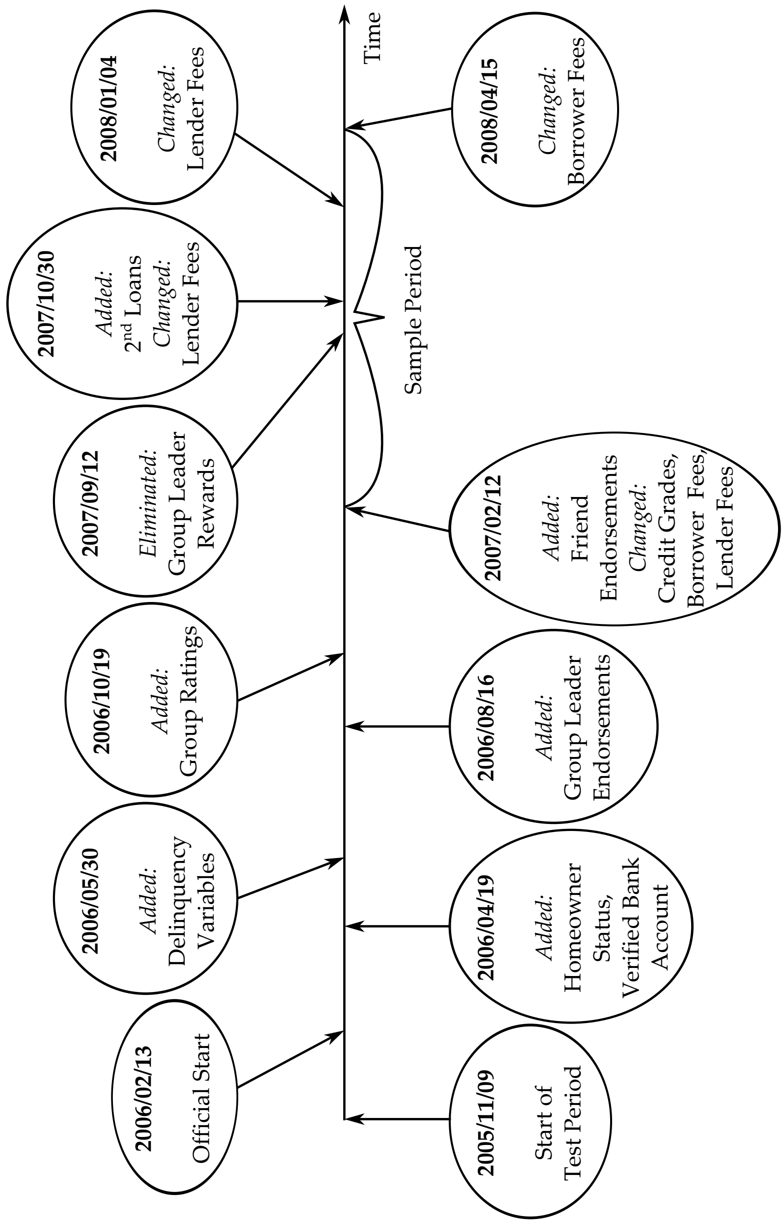


Figure 4.3: Important Policy Changes on Prosper.com

Table 4.8: Variable Definitions

Variable	**Definition**
Group Leader Bid	The group leader places a bid on the listing.
Group Leader Endorsement	The group leader writes an endorsement (a short text statement) on the borrower / her listing (before the loan is funded or the listing expires).
Only GL Bid	The group leader places a bid on the listing but does not write an endorsement.
Only GL Endorsement	The group leader writes an endorsement for the listing but does not place a bid.
GL Bid & GL Endorsement	The group leader places a bid on the listing and writes an endorsement.
Before / After	The listing is created before / after the elimination of group leader rewards. "Before" is the reference in the multivariate analyses.
No Group	The listing is not posted in any group.
No-Reward (Group) / Reward (Group)	If the group leader does not request a reward for any listing posted in the group in the sample period, the group is considered as a no-reward group. Otherwise the group is considered as a reward group. "No-Reward Group" is the reference in the multivariate analyses.
Vetting	The group leader asks the borrower to provide information.
Listing Review Requirement	The group leader reviews the listing before it is open for bidding by the lenders.
Group Leader Offers Help	The group leader provides help in designing and writing the listing.
Credit Grade: AA/A, B, C, D, E, HR	Each borrower is assigned a credit grade based on her Experian credit score. AA designates the lowest risk, HR the highest. "Credit Grade: AA/A" is the reference in the multivariate analyses.
Debt-to-Income Ratio	The debt-to-income ratio of the borrower at the time the listing was created. This value is capped at 1.01.
Is Borrower Home Owner	Specifies whether or not the member is a verified homeowner at the time the listing is created.
Income Information Unavailable / $1-24,999 / $25,000-49,999 / $50,000-74,999 / $75,000-99,999 / $100,000+	The income range of the borrower at the time the listing is created. "Income Information Unavailable" is the reference in the multivariate analyses.
Full-Time / Part-Time / Self-Employed / Retired / Not Employed	The occupation status of the borrower at the time the listing is created. "Full-Time" is the reference in the multivariate analyses.
Current Delinquencies	Number of current delinquencies at the time the listing is created.
Delinquencies Last 7 Years	Number of delinquencies in the last 7 years at the time the listing is created.
Public Records Last 10 Years	Number of public records in the last 10 years at the time the listing is created.
Total Credit Lines	Number of total credit lines at the time the listing is created.
Inquiries Last 6 Months	Number of inquires in the last 6 months at the time the listing is created.
Amount Delinquent (in $1,000)	The monetary amount delinquent at the time this listing is created. (in $1,000)
Public Records Last 12 Months	Number of public records in the last 12 months at the time the listing is created.
Current Credit Lines	Number of current credit lines at the time the listing is created.
Open Credit Lines	Number of open credit lines at the time the listing is created.
Revolving Credit Balance (in $1,000)	The monetary amount of revolving credit balance at the time the listing is created. (in $1,000)
Bankcard Utilization	Describes whether the borrower uses a banking card for her transactions.
Length Status Months	The length in months of the employment status of the borrower at the time the listing is created.
Amount Requested (in $1,000)	The amount requested by the borrower in the listing. (in $1,000)
Duration	The time for which the listing is open for bidding by potential lenders.
Listing Closed As Soon As Funded	The listing is automatically closed as soon as it is completely funded, i.e. once the total amount bid reaches or exceeds the amount requested.

5 Concluding Remarks

In this dissertation, I have analyzed various aspects of two-sided markets with a particular focus on the importance of online platforms. As shown throughout the dissertation, the inception of such platforms results in challenges and implications for the users they attract, for users of the traditional offline platforms and in particular for platform intermediaries.

As an example, in the first paper of this dissertation I have shown that the creation of companion websites to traditional consumer magazines results in a crowding-out effect of platform revenue from traditional offline reading and advertising. The reason is that both readers and advertisers substitute online for offline. The model I develop for this analysis is not limited to the case of consumer magazines and their companion websites, but can be potentially modified to analyze substitution effects in other two-sided markets.

As the Internet grows further, new business models and types of market platforms are created. Therefore, it is increasingly important for both platform operators and policymakers to quantify correctly the network effects that arise on such platforms. One innovative approach to this problem I have provided in the second paper where I have derived a sufficient (and under mild conditions also necessary) test for network effects in potentially two-sided monopoly platforms.

A different challenge is analyzed in the third paper of this dissertation. It shows how financial disintermediation on innovative online social lending platforms may lead to adverse incentives for platform members if the reward scheme is not set up carefully. The results from this experiment provide important implications for the question of how retail consumers can be protected against unscrupulous lending and thus the ongoing debate about the proper regulatory framework for consumer lending.

Together, the three papers of this dissertation provide help to researchers who want to improve their understanding how two-sided markets function and what their partiularities are – especially with respect to the rising importance of the Internet.

Bibliography

Ackerberg, D. A. and Gowrisankaran, G. (2006). Quantifying Equilibrium Network Externalities in the ACH Banking Industry. *RAND Journal of Economics*, 37(3):738–761.

Akerlof, G. A. (1970). The Market for “Lemons”: Quality Uncertainty and the Market Mechanism. *Quarterly Journal of Economics*, 84(3):488–500.

Anderson, S. P. and Coate, S. (2005). Market Provision of Broadcasting: A Welfare Analysis. *Review of Economic Studies*, 72(4):947–972.

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

Armstrong, M. (1998). Network Interconnection in Telecommunications. *Economic Journal*, 108(448):545–564.

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

Berger, A. N., Miller, N. H., Petersen, M. A., Rajan, R. G., and Stein, J. C. (2005). Does function follow organizational form? Evidence from the lending practices of large and small banks. *Journal of Financial Economics*, 76(2):237–269.

Bergstresser, D., Chalmers, J. M. R., and Tufano, P. (2009). Assessing the Costs and Benefits of Brokers in the Mutual Fund Industry. *Review of Financial Studies*, 22(10):4129–4156.

Berry, S., Levinsohn, J., and Pakes, A. (1995). Automobile Prices in Market Equilibrium. *Econometrica*, 63(4):841–890.

Berry, S. T. (1994). Estimating Discrete-Choice Models of Product Differentiation. *RAND Journal of Economics*, 25(2):242–262.

Bolton, P., Freixas, X., and Shapiro, J. (2007). Conflicts of interest, information provision, and competition in the financial services industry. *Journal of Financial Economics*, 85(2):297–330.

Bolton, P. and Scharfstein, D. S. (1996). Optimal Debt Structure and the Number of Creditors. *Journal of Political Economy*, 104(1):1–25.

Bresnahan, T. F. (1987). Competition and Collusion in the American Automobile Industry: The 1955 Price War. *Journal of Industrial Economics*, 35(4):457–482.

Brynjolfsson, E. and Kemerer, C. F. (1996). Network Externalities in Microcomputer Software: An Econometric Analysis of the Spreadsheet Market. *Management Science*, 42(12):1627–1647.

Caillaud, B. and Jullien, B. (2001). Competing cybermediaries. *European Economic Review*, 45(4-6):797–808.

Caillaud, B. and Jullien, B. (2003). Chicken & egg: competition among intermediation service providers. *RAND Journal of Economics*, 34(2):309–328.

Church, J. and Gandal, N. (2006). Platform Competition in Telecommunications. In Majumdar, S. K., Vogelsang, I., and Cave, M., editors, *Handbook of Telecommunications Economics*, volume 2, pages 119–155. Elsevier.

Circle of Online Marketers (2010). OVK Online-Report 2010/01 - Zahlen und Trends im Überblick. Technical report, Circle of Online Marketers (OVK).

Dranove, D. and Gandal, N. (2003). The DVD-vs.-DIVX Standard War: Empirical Evidence of Network Effects and Preannouncement Effects. *Journal of Economics & Management Strategy*, 12(3):363–386.

Economides, N. (1996). The economics of networks. *International Journal of Industrial Organization*, 14(6):673–699.

Economides, N. and Himmelberg, C. (1995). Critical Mass and Network Size with Application to the US Fax Market. Working Paper 95-11, New York University, Leonard N. Stern School of Business, Department of Economics.

Economides, N. and Katsamakas, E. (2006). Two-Sided Competition of Proprietary vs. Open Source Technology Platforms and the Implications for the Software Industry. *Management Science*, 52(7):1057–1071.

Eisenmann, T., Parker, G., and van Alstyne, M. W. (2006). Strategies for Two-Sided Markets. *Harvard Business Review*, 1463:1–12.

Ellison, G., Fudenberg, D., and Möbius, M. (2004). Competing Auctions. *Journal of the European Economic Association*, 2(1):30–66.

Evans, D. S. (2003). The Antitrust Economics Of Multi-Sided Platform Markets. *Yale Journal on Regulation*, 20(2):325–381.

Evans, D. S., Hagiu, A., and Schmalensee, R. (2005). A Survey of the Economic Role of Software Platforms in Computer-based Industries. *CESifo Economic Studies*, 51(2-3):189–224.

Feenstra, R. C. and Levinsohn, J. A. (1995). Estimating Markups and Market Conduct with Multidimensional Product Attributes. *Review of Economic Studies*, 62(1):19–52.

Ferrando, J., Gabszewicz, J. J., Laussel, D., and Sonnac, N. (2003). Two-Sided Network Effects and Competition: An Application to Media Industries. Working paper, CREST-LEI and EUREQUA, CORE, Université Catholique de Louvain, GREQAM, Université de la Méditerranée, CREST-LEI and Université de Paris II.

Filistrucchi, L. (2005). The Impact of Internet on the Market for Daily Newspapers in Italy. Working Paper ECO2005/12, European University Institute.

Gabszewicz, J. J., Laussel, D., and Sonnac, N. (2001). Press advertising and the ascent of the 'Pensée Unique'. *European Economic Review*, 45(4-6):641–651.

Gans, J. S. and King, S. P. (2003). The Neutrality of Interchange Fees in Payment Systems. *Topics in Economic Analysis & Policy*, 3(1):1069 ff.

Gentzkow, M. (2007). Valuing New Goods in a Model with Complementarity: Online Newspapers. *American Economic Review*, 97(3):713–744.

Gowrisankaran, G. and Stavins, J. (2004). Network Externalities and Technology Adoption: Lessons from Electronic Payments. *RAND Journal of Economics*, 35(2):260–276.

Grajek, M. (2004). *Network Effects, Compatibility, and Adoption of Standards: Essays in Empirical Industrial Economics.* Dissertation, Humboldt-Universität zu Berlin, School of Business and Economics.

Hagiu, A. (2007). Multi-Sided Platforms: From Microfoundations To Design And Expansion Strategies. Working Paper 07-094, Harvard Business School.

Hermalin, B. E. and Katz, M. L. (2006). Your Network or Mine? The Economics of Routing Rules. *RAND Journal of Economics*, 37(3):692–719.

Holmström, B. (1979). Moral Hazard and Observability. *Bell Journal of Economics*, 10(1):74–91.

Holmström, B. and Tirole, J. (1997). Financial Intermediation, Loanable Funds, and the Real Sector. *Quarterly Journal of Economics*, 112(3):663–691.

Hulme, M. K. and Wright, C. (2006). Internet Based Social Lending: Past, Present and Future. Technical report, Social Futures Observatory.

Inderst, R. and Ottaviani, M. (2009). Misselling through Agents. *American Economic Review*, 99(3):883–908.

IVW (2009). Anlage 1 zu den IVW-Richtlinien für Online-Angebote (Version 2.2). Technical report, Informationsgemeinschaft zur Feststellung der Verbreitung von Werbeträgern e.V. (IVW).

Iyer, R., Khwaja, A. I., Luttmer, E. F. P., and Shue, K. (2009). Screening in New Credit Markets: Can Individual Lenders Infer Borrower Creditworthiness in Peer-to-Peer Lending? Working paper, MIT Sloan School of Management.

Jeon, D.-S., Laffont, J.-J., and Tirole, J. (2004). On the "Receiver-Pays" Principle. *RAND Journal of Economics*, 35(1):85–110.

Judge, G. G., Hill, R. C., Griffiths, W. E., Lütkepohl, H., and Lee, T.-C. (1988). *Introduction to the Theory and Practice of Econometrics*. John Wiley & Sons, Inc., 2nd edition.

Jullien, B. (2001). Multi-Sided Markets: Competing with Network Externalities and Price Discrimination. Working paper, Centre for Economic Research Policy.

Jullien, B. (2005). Two-sided Markets and Electronic Intermediaries. *CESifo Economic Studies*, 51(2-3):233–260.

Kaiser, U. (2002). Optimal Cover Prices and the Effects of Website Provision on Advertising and Magazine Demand. ZEW Discussion Papers 02-54, ZEW - Zentrum für Europäische Wirtschaftsforschung / Center for European Economic Research.

Kaiser, U. (2006). Magazines and their Companion Websites: Competing Outlet Channels? *Review of Marketing Science*, 4(1):1–34.

Kaiser, U. and Song, M. (2009). Do media consumers really dislike advertising? An empirical assessment of the role of advertising in print media markets. *International Journal of Industrial Organization*, 27(2):292–301.

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

Katsamakas, E. and Bakos, Y. (2004). Design and ownership of two-sided networks: implications for Internet intermediaries. Doctoral Dissertation Paper 1, Department of Information, Operations and Management Science, Stern School of Business, New York University.

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

Katz, M. L. and Shapiro, C. (1994). Systems Competition and Network Effects. *Journal of Economic Perspectives*, 8(2):93–115.

Kind, H. J., Nilssen, T., and SÃ¸rgard, L. (2005). Advertising on TV: Under- or Over-provision? Memorandum 15/2005, University of Oslo, Department of Economics, University of Oslo, Department of Economics, P. O.Box 1095 Blindern, N-0317 Oslo, Norway.

Liebowitz, S. J. and Margolis, S. E. (1994). Network Externality: An Uncommon Tragedy. *Journal of Economic Perspectives*, 8(2):133–150.

Lin, M., Prabhala, N. R., and Viswanathan, S. (2009). Judging Borrowers By The Company They Keep: Social Networks and Adverse Selection in Online Peer-to-Peer Lending. Working paper, University of Maryland, College Park.

Panzar, J. C. and Rosse, J. N. (1987). Testing For “Monopoly” Equilibrium. *Journal of Industrial Economics*, 35(4):443–456.

Parker, P. M. and Röller, L.-H. (1997). Collusive Conduct in Duopolies: Multimarket Contact and Cross-Ownership in the Mobile Telephone Industry. *RAND Journal of Economics*, 28(2):304–322.

Pope, D. G. and Sydnor, J. R. (forthcoming). What’s in a Picture? Evidence of Discrimination from Prosper.com. *Journal of Human Resources.*

Ravina, E. (2008). Love & Loans: The Effect of Beauty and Personal Characteristics in Credit Markets. Working paper, New York University.

Reisinger, M. (2004). Two-Sided Markets with Negative Externalities. Discussion Paper in Economics 478, University of Munich, Department of Economics.

Reiss, P. C. and Wolak, F. A. (2007). Structural Econometric Modeling: Rationales and Examples from Industrial Organization. In Heckman, J. J. and Leamer, E. E., editors, *Handbook of Econometrics*, volume 6, Part 1, pages 4277–4415. Elsevier.

Rochet, J.-C. and Tirole, J. (2002). Cooperation among Competitors: Some Economics of Payment Card Associations. *RAND Journal of Economics*, 33(4):549–570.

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

Rochet, J.-C. and Tirole, J. (2006). Two-Sided Markets: A Progress Report. *RAND Journal of Economics*, 37(3):645–667.

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

Roson, R. (2005). Two-Sided Markets: A Tentative Survey. *Review of Network Economics*, 4(2):142–160.

Rysman, M. (2004). Competition Between Networks: A Study of the Market for Yellow Pages. *Review of Economic Studies*, 71(2):483–512.

Rysman, M. (2007). An Empirical Analysis of Payment Card Usage. *Journal of Industrial Economics*, 55(1):1–36.

Schmalensee, R. (2002). Payment Systems and Interchange Fees. *Journal of Industrial Economics*, 50(2):103–122.

Shy, O. (2001). *The Economics of Network Industries.* Cambridge University Press.

Stein, J. C. (2002). Information Production and Capital Allocation: Decentralized versus Hierarchical Firms. *Journal of Finance*, 57(5):1891–1921.

Stiglitz, J. E. and Weiss, A. (1981). Credit Rationing in Markets with Imperfect Information. *American Economic Review*, 71(3):393–410.

Sufi, A. (2007). Information Asymmetry and Financing Arrangements: Evidence from Syndicated Loans. *Journal of Finance*, 62(2):629–668.

The White House (2009). Remarks by the President on 21st century financial regulatory reform. *Office of the Press Secretary*.

Wright, J. (2002). Access Pricing under Competition: An Application to Cellular Networks. *Journal of Industrial Economics*, 50(3):289–315.

Wright, J. (2003). Optimal card payment systems. *European Economic Review*, 47:587–612.

Wright, J. (2004). One-sided Logic in Two-sided Markets. *Review of Network Economics*, 3(1):44–64.

List of Figures

List of Tables

Selbständigkeitserklärung

Ich bezeuge durch meine Unterschrift, dass meine Angaben über die bei der Abfassung meiner Dissertation benutzten Hilfsmittel, über die mir zuteil gewordene Hilfe sowie über frühere Begutachtungen meiner Dissertation in jeder Hinsicht der Wahrheit entsprechen.

Berlin, den 14.03.2011