Discussion Paper No. 05-52

The Virtual Location of E-Tailers

Evidence from a B2C E-Commerce Market

Julia Häring

Discussion Paper No. 05-52

The Virtual Location of E-Tailers

Evidence from a B2C E-Commerce Market

Julia Häring

Non-technical Summary

Today, it is widely accepted that the Internet is beneficial for competition among retailers, but that these effects are not as pervasive as it had been expected. One factor mitigating competition is the fact that online retailers (e-tailers) are not only a single mouse-click apart from each other, which renders price comparisons difficult. It should be obvious from everyday observation that not all online retailers can be found by uninformed consumers in an equally easy way. Instead, some online shops are easier to be found than others. This seems to be an analogy to the location in the physical world where shop owners invest considerable amounts of money in superior locations, for example in highly frequented shopping malls or pedestrian areas. On the Internet, search engines or news portals are the highly frequented locations where hyperlinks leading to other sites on the Web are noticed by Internet surfers with a larger probability. This Internet analogy to the concept of location in the physical is termed *virtual location.*

In this paper, the determinants of the virtual location of e-tailers are analysed empirically. As most consumers seem to use relatively little sophisticated techniques when searching the Web, an outstanding virtual location is a crucial factor for the attraction of potential customers to e-commerce sites. Thus, optimising their virtual location becomes an integral part of e-tailers' overall strategy. In the empirical analysis, the virtual location will be proxied by an online retailer's position in the Google search results list and by online advertising activities, since these may be the most important factors making it more likely for consumers to get to know about a specific online shop. The observed activities of online advertising are context dependent banner ads and sponsored links in the 10 most-widely used Internet search engines.

The results suggest that it is optimal for e-tailers to complement a high search engine rank with investments in online advertising. Moreover, banner ads seem to serve as price advertising mechanism, whereas sponsored links rather seem to be used in order to signal outstanding customer service.

The Virtual Location of E-Tailers
Evidence from a B2C E-Commerce Market

JULIA HÄRING

ZEW (Centre for European Economic Research)

April 2005

Abstract
An Internet retailer's (e-tailer's) outstanding virtual location enhances the probability of being noticed by potential customers. The notion of a virtual location for e-tailers refers to the analogy to the physical location. In the empirical analysis, an e-tailer's Internet search engine rank as well as its advertising activities in search engines serve as proxies for the virtual location. The results suggest that it is optimal for e-tailers to complement a high search engine rank with investments in online advertising. Moreover, banner ads seem to serve as price advertising mechanism, whereas sponsored links rather seem to be used in order to signal outstanding customer service.

Address: Centre for European Economic Research (ZEW)
Research Group Information and Communication Technologies
P.O. Box 10 34 43
D–68034 Mannheim
Germany

Phone: +49/621/1235–134

Fax: +49/621/1235–333

E-Mail: haering@zew.de

1 Introduction

Today, it is widely accepted that the Internet is not the "great equalizer" it was expected to be for competition and retailer prices (see for example Smith and Brynjolfsson, 2001; Clay, Krishnan, Wolff and Fernandes, 2002). Empirical analyses have mainly focused on differentiation in retailer service as a reason for price premiums, which induce price dispersion and impede the observation of the law of one price. Smith, Bailey and Brynjolfsson (2000, p. 110) argue that there are certain general interest websites which make some online retailers easier to be found than their competitors and refer to the huge amounts of money invested in portals and "content sites". This phenomenon is interpreted by the authors as "neural real estate" (see Smith et al., 2000, p. 110).

It should be obvious from everyday observation that not all online retailers can be found by uninformed consumers in an equally easy way. Instead, some online shops are easier to be found than others. This seems to be an analogy to the location in the physical world where shop owners invest considerable amounts of money in superior locations, for example in highly frequented shopping malls or pedestrian areas. On the Internet, search engines or news portals are the highly frequented locations where hyperlinks leading to other sites on the Web are noticed by Internet surfers with a larger probability. This Internet analogy to the concept of location in the physical is termed *virtual location.* This follows the view by Hunter (2003) who argues from a juridical point of view that cyberspace is perceived as a place with spatial characteristics, and discusses the implications for legal regulation.

This paper analyses empirically the determinants of the virtual location of e-tailers[1]. As most consumers seem to use relatively little sophisticated techniques when searching the Web (this is one of the results of the study by Machill, Neuberger, Schweiger and Wirth, 2003), an outstanding virtual location is a crucial factor for the attraction of potential customers to e-commerce sites. Thus, optimising their virtual location becomes an integral part of e-tailers' overall strategy. In the empirical analysis, the virtual location will be proxied by an online retailer's position in the Google search results list and by online advertising activities, since these may be the most important factors making it more likely for consumers to get to know about a specific online shop.[2] The observed activities of online advertising are context dependent banner ads and sponsored links in the 10 most-widely used Internet search engines.[3]

A recent study investigating the market for Internet search engines in German language by Machill et al. (2003) suggests that having a prominent location on the Internet is crucial for attracting consumers to websites.[4] The study consists of two parts: a survey among operators of search engines in German language, and an

---

[1]Throughout the paper, the terms 'e-tailer' and 'online retailer' are used interchangeably and refer to firms selling products online via e-commerce websites.

[2]There will of course be repeat purchasers or potential customers who directly access a specific e-tailer which they are aware of due to retailer branding or word of mouth. Hence, similarly to the physical location, the virtual location is crucial for the attraction of new customers.

[3]Examples for banner ads and sponsored links are depicted in Figures 2 and 3 in the Appendix.

[4]The author of this paper is not aware of any comparable study for a further language area.

experiment, in which the steps of Internet users when solving specified search tasks are observed and analysed. Machill et al. (2003, p. 92) report extensive linkages between different search engines, as many of them share the same search technology. Furthermore, it is well-known that several highly-frequented search engines rely on the Google technology, and that the Google website itself has a dominant market share. Both factors combined lead to consumers being highly dependent on the information provided by just a few search sites on the Web.

Most consumers are neither aware of the economic dominance of the Google technology, nor aware of the existence of sponsored links on results pages (see Machill et al., 2003, p. 94).[5] Furthermore, search engine providers state in the survey, that users click-through to just a few links of the results pages and that nearly 70% of the users do not examine more than the first two pages of results, corresponding to the first 20 hits. This is confirmed by the experimental results of the study (see Machill et al., 2003, p. 255): 81% of the participants evaluated only the first page of results, further 13% the first and second pages, implying that only 6% considered more than the first 20 entries. The results suggest that being visible on the first or second page of the results list should improve the ability to attract consumers quite a lot. A means to reach this goal is the use of advertising in Internet search engines, thus serving as a virtual location. The virtual location being proxied by both the rank in the Google list of results and contextual advertising in different Internet search engines is in line with these arguments.

The contribution of this paper is twofold: This is – to the best of the author's knowledge – the first attempt to relate the virtual location to e-tailer strategy resulting from profit-maximising behaviour. Second, the analysis is based on a unique set of primary data from the online market for contact lenses relevant for consumers living in Germany. The data set contains observations for 146 different online retailers collected between March and September 2002 on a monthly basis. It is not a selected sample but it represents the whole population of online shops, which are relevant for consumers living in Germany.

The primary data set contains the prices and range of products offered by retailers in approximately the whole market for contact lenses. The data are merged with retailer characteristics as well as information on the virtual location of online retailers. Neither is the observed product range restricted to a predetermined subset of products, nor are the retailers selected. The observations are collected directly at the e-tailers' websites instead of using shopbot data or data from price comparison websites.

The results presented in this paper suggest that it is optimal for profit maximising e-tailers to complement a high search engine rank by investments in online advertising. Moreover, banner ads seem to serve as price advertising mechanism, whereas sponsored links rather seem to be used by e-tailers in order to signal out-

---

[5]This is supported by the fact that in contrast to journalism, where advertising and editorial contents are usually clearly separated according to a code of conduct, this is not common in the field of Internet search engines (see Machill et al., 2003, p. 92). Also Silk, Klein and Berndt (2001, p. 140) report a "blurring of the traditional distinction between advertising and editorial content on the Internet."

standing customer service. The search engine rank appears to remain relatively stable over time suggesting that the virtual locations for the whole market are in an equilibrium during the period under observation.

The paper is organised as follows: Section 2 describes the related literature, and Section 3 provides background information on the market for online advertising. Section 4 describes the data set, which is analysed in Section 5. Section 6 concludes.

2 Related Literature

So far, the existing literature on aspects of the virtual location has been sparse. The growing importance for online retailers to be prominently placed on the Internet has been brought to the discussion by Smith et al. (2000). Hunter (2003) discusses the metaphor of "cyberspace as place". Based on a theoretical model, Smith (2002) distinguishes between online retailers with a high awareness on the consumer side and those with a low awareness and focuses on implications on the pricing strategies of the two groups. That location merges with brand in the virtual space, is an observation by Tang and Lu (2001) who analyse price dispersion among online retailers.

There is a growing strand of literature dealing with advertising on the Internet, most of which tackles the effect of price advertising on competition (see for example Stahl (2000) or Baye and Morgan (2001)). Stahl (2000) focuses on the relation between online price advertising and pricing in e-commerce and the welfare implications. Baye and Morgan (2001) analyse e-tailers' incentives to advertise prices on a gatekeeper's site (such as a price comparison site where firms pay in order to have their prices listed) and the competitive effects of such price advertising. The predictions of this model are tested in an experimental setting by Morgan, Orzen and Sefton (2003). From a theoretical point of view, Baye and Morgan (2004) consider the effects of informational and promotional advertising in a common model.[6] Similarly to the other studies, the effects of both types of advertising on price competition are investigated.

There are two main points differing between the aforementioned empirical papers and this paper. First, other studies mostly rely on data from a specific gatekeeper's website (for example Baye, Morgan and Scholten (2004)). An important drawback for such approaches is the neglect of retailers never listing on the specific price comparison site considered which may imply a highly selective sample. In contrast to this, also retailers never advertising are observed in the empirical analysis presented in Section 5. A second difference between the papers cited above and this paper is their focus on price advertising and price competition, whereas this paper deals with e-tailer *brand* advertising as a proxy for the virtual location of e-tailers.

The role of location on the Internet has received attention in a recent paper from Baye, Gatti, Kattuman and Morgan (2004), where a different aspect of virtual

---

[6]Informational advertising refers to the listing of prices on a price comparison site attracting price sensitive consumers, whereas promotional advertising is understood as creating loyal customers and thus enticing website traffic away from price comparison sites.

4

location is investigated: effects of the position of e-tailers in the price quotation list of an Internet price comparison site on the decision to click-through to a specific e-tailer are analysed. The estimates suggest that customers favour products that are listed higher on the screen. It is important to note that this effect is modelled and estimated independently of the relative price compared to the other firms in the list.[7] Also Ellison and Fisher Ellison (2004) include the rank of the retailer in their estimation explaining the demand for computer memory modules, but retailers are automatically sorted according to price on the shopbot site underlying their sample. Therefore, the effect of the order cannot be distinguished from the price effect. Smith and Brynjolfsson (2001) find in an empirical analysis of the click-through behaviour at an Internet shopbot for books that a considerable fraction of consumers does not decide in favour of the cheapest retailer. The influence of an e-tailer's position on the screen can however not be analysed, as retailers are automatically sorted according to price.

The question of how to identify the ideal location for paid online advertising is addressed by Bhatnagar and Papatla (2001). They directly address the issue of narrowly targeting potential customers on the Internet. Their discussion starts from the consideration that a firm's customer segment would be best found on the websites of its competitors which however would deny to sell advertising space. The discussion focuses on ways to identify adequate websites to advertise at and considers consumer search behaviour in the analysis. Consumer response to banner ads is discussed and empirically analysed in Chatterjee, Hoffman and Novak (2003) or Manchanda, Dubé, Goh and Chintagunta (2004), for example.

In this paper, it is argued that the virtual location is tied to e-tailers' rankings and context dependent advertising efforts in Internet search engines. This view is supported by Machill et al. (2003) where a systematic evaluation of the role and power of Internet search engines can be found. The study discusses the role and market power of search engines in the German language area of the Internet but the results should be transferable to the English part of the Internet without major obstacles. The authors focus on the market structure in the search engine market and additionally present an extensive experimental study of user behaviour when using Internet search engines (selected results of this study have already been reported in the Introduction).

3   The Market for Online Advertising

Online advertising faces increasing problems of lacking acceptance by Internet users. A recent study for Germany revealed that the proportion of Internet users tolerating advertising on websites as a necessary instrument for financing websites has shrunk to 41 percent from 53 percent at the beginning of 2001 (see Fittkau & Maaß (2003), cited in ECIN (2003a)). Simultaneously, the proportion of Internet users claiming to ignore online advertising has risen from 34 to 41 percent. Both Kenny and Marshall

---

[7]**?** are able to assess the roles of the price and the position on the screen separately, as firms are not listed according to price on the price comparison site they observe.

(2000) and Hoffman and Novak (2000) report average click rates as low as 0.5 percent for banner ads. Hoffman and Novak (2000) argue in this context that the optimal placement of online advertising activities is crucial in order to achieve higher click rates.

One way to an exact targeting of the relevant consumer group is contextual marketing (see Luo and Seyedian, 2004). Contextual marketing is defined by Kenny and Marshall (2000, p. 120) as using the Internet "to deliver tailored messages and information to customers at the point of need". One instrument of contextual marketing are banner ads and sponsored links appearing together with the results list after search engine queries for specific keywords, which have been chosen by the marketer. *Banner ads* refer to ads which are graphically emphasised, for example by using coloured boxed, graphics or pictures to draw attention on them. In contrast, *sponsored links* have an impact by their unconspicious placement on top of the results list. The optical appearance of these prominently placed links is meant to differ as least as possible from the ordinary list of results. Examples for banner ads and sponsored links are depicted in Figures 2 and 3 in the Appendix.

According to Jupiter Media Metrix (as cited in ECIN (2003b)), 42 percent of online purchases are initiated via a preceding search which makes context dependent advertising in Internet search engines an ideal advertising channel for online retailers. If advertising is linked to specified search terms, people interested in the product qualify as possible customers by the word(s) they actively search for. In addition to the better targeting of the audience, the advertiser incurs costs only if the user actually clicks on the banner ad or sponsored link.

The evolution of online advertising in the market for glasses and contact lenses during the period covered by the data set underlying this paper is depicted in Figure 1.[8] According to W3SCAN.COM, the spending for online advertising rose from 29,879 Euro during the first wave of data collection in March 2002 to 48,855 Euro during the weeks following the last collection of data in September 2002. In total, 614,063 Euro were spent for online advertising for optics in 2002. These figures are published online by the company W3SCAN.COM on their website (www.w3scan.com).

4   Data

The analysis of markets for optometric devices has a long tradition in economics (see for example Benham, 1972; Kwoka, 1984). The online market for contact lenses was chosen due to several criteria which make the products suitable for both being sold via e-commerce and being analysed empirically.[9] The data set used for the analysis is condensed from a data set with monthly observations of e-tailers for contact lenses which were observed between March and September 2002.

---

[8]Unfortunately, there is no separate information on the segment for contact lenses excluding glasses available.

[9]These criteria are discussed in Häring (2003) where also a more detailed description of the data set can be found.

Figure 1: Market for online advertising in the weeks of data collection
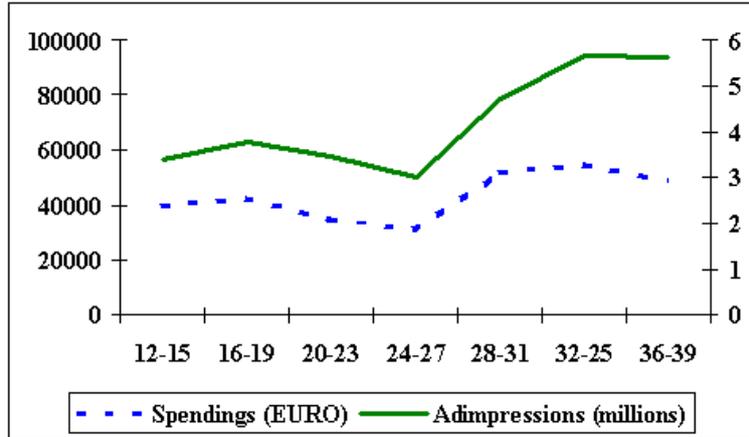
Source: www.w3scan.com

Before collecting the primary data set, the online shops for contact lenses had to be identified. This was done by searching for the German word for contact lenses in its two possible spellings (*Kontaktlinsen* and *Contactlinsen*) in the ten most-widely used Internet search engines at the beginning of March 2002. The list of search engines can be found in Table 5 in the Appendix. From each of these search queries, the first 250 results were evaluated in order to identify sellers of contact lenses.

The primary data set contains monthly information on both the range of products and the prices offered by the e-tailers which results in more than 20,000 price observations. Relevant product attributes of the contact lenses and the service characteristics of the online shops were evaluated once during the period of data collection and then merged with the price data. This was appropriate since none of the online shops underwent major changes, and also no product was relaunched with retention of its original name.

Furthermore, information on the e-tailers' virtual locations was included. The virtual location is proxied by context dependent banner ads and sponsored links in the ten most-widely used search engines[10] and by the rank in the Google search engine at each time of data collection when a search for the German word for contact lenses in its two possible spellings was conducted. Both forms of online advertising are linked to the searches for contact lenses in the search engines as described above. The ranks in the Google list of results were only considered for the first 10 pages of results (i.e. ranks 1 to 100). Price comparison sites played no role in the market for contact lenses at the time of data collection.

For the analysis of the link between e-tailer strategies and their virtual location, the original data set was condensed. The characteristics of the online shops underwent a factor analysis in order to obtain factors describing the service of the online retailers. Names were assigned to the factors according to the underlying variables they represent. The result are five factors indicating: Convenient navi-

---

[10]A list of these can be found in Table 5 in the Appendix.

gation, superior customer service, a favourable return policy & supply of lens care products, security and trustworthiness features of the websites, and special services for customers using contact lenses for the first time.[11] The width of the product range offered by the e-tailers is measured by the number of different products offered. Clearly, the probability of the online shop covering different product segments of the contact lense market increases with the number of distinct products offered. Each e-tailer's overall price level in comparison to its competitors is proxied by the average of its standardised prices. For this standardisation, each product price is divided by the average price over all e-tailers for this product.

The result is an unbalanced panel data set with 929 observations on a monthly basis for 146 different e-tailers. On average, the e-tailers are observed in 6.4 of the seven months of data collection. Descriptive statistics of the variables used in the empirical analysis can be found in the Appendix.

5 Empirical Analysis

5.1 Framework

The determinants of virtual location are analysed using data from the online market for contact lenses. In the case of online advertising, it is obvious that firms deliberately choose whether or not they advertise and which kind(s) of advertising to invest in. But also the rank in the Google results list can be interpreted as an outcome of the profit maximising behaviour of e-tailers. Thus, both dimensions reflect the underlying latent profit maximisation of the e-tailers, and both dimensions of the virtual location can be investigated using the latent profit index framework which will briefly be described in the following.

The latent profit is an analogue to the latent utility in the consumer choice literature (see McFadden, 1974). The latent profit index of an e-tailer when choosing alternative j out of J possible choices is not observable and assumed to consist of a systematic and a stochastic component:

$$y^* = \mathbf{x}\beta + \epsilon \tag{1}$$

where the latent profit y^* is not observable, $\mathbf{x}$ is a vector of observable characteristics, β is a parameter vector, and ϵ is a stochastic error term. The observed outcome y takes on one of the values 1, ..., J indicating the chosen alternative. Thereby, some information about the latent index is revealed but the underlying profit level cannot be fully recovered. In order to use the observed information on the virtual location of e-tailers, assumptions about the decision process are made in order to estimate the relationship between various e-tailer attributes and their virtual location. These are explained in the context of the estimation problem in the following two subsections.

---

[11]The construction of these factors is explained in Häring (2003).

5.2 An Empirical Analysis of The Virtual Location

5.2.1 Correlation between Search Engine Rank and E-Tailer Strategy

It has been argued in the previous sections that a superior position in the results list of search engines is of considerable importance. Since most of the popular search engines nowadays rely on the Google technology, the achieved position in the Google list of results is a valid proxy for the virtual location of online retailers.

Regarding the result of Machill et al. (2003) according to which most consumers evaluate only the first or the first two pages of results when using Internet search engines, the observed Google ranks are grouped into categories for the first (position 1-10) and the second (position 11-20) page of results in the list. The rest of the e-tailers is sorted into the third category (position 21 and above). The frequency distribution over these ordered categories is denoted in Table 6 in the Appendix.

In order to identify the strategies pursued by e-tailers which are correlated with a high Google rank, the rank is explained by the width of the product range the e-tailer offers, its relative price level, the amount of online advertising and the five factors describing retailer services. Since the rank categories are ordered (a higher rank is better), an ordered discrete choice model is the appropriate tool for analysis. The latent index y^* in Equation 1 can be used to estimate the correlations between the e-tailers' strategies and their Google rank using an ordered probit model[12]. In the case of three categories it is assumed that we observe:

$$\begin{aligned}
y &= 0 & &\text{if } y^* \leq \mu_1 \\
y &= 1 & &\text{if } \mu_1 < y^* \leq \mu_2 \\
y &= 2 & &\text{if } \mu_2 < y^*
\end{aligned}$$

where $\mu_1 < \mu_2$ are unknown threshold parameters. By using the ordered probit model, a standard normal distribution for the error term $\epsilon \mid \mathbf{x}$ is assumed, and the unknown parameters β, μ_1 and μ_2 are estimated by maximum likelihood. The estimated parameter value $\hat{\beta}_k$ for regressor x_k does not correspond to the marginal effect of a change in x_k on the conditional probability to observe outcome j, denoted as $Pr(y = j \mid \mathbf{x})$. These marginal effects $\partial p_j(\mathbf{x})/\partial x_k$ have to be computed separately for each outcome category. This has been done in Table 1 where the estimation results can be found. The standard errors are estimated by nonparametric bootstrapping in order to account for the inclusion of generated regressors, which were obtained by the factor analysis.

The estimated marginal effects reveal that a wider product range and a higher price level are associated to a higher probability of being ranked on the first two pages of results in the Google search engine. Moreover, the probability of achieving a rank between 1 and 20 decreases with the service level offered by the e-tailers. But poor service can apparently be offset by investments in online advertising. The probabilities of being ranked on the first two pages of results are positively correlated with the numbers of banner ads and of sponsored links, respectively. On the contrary, the probability of being ranked lower than the first 20 entries increases

---

[12]For a textbook treatment of the ordered probit model, see Wooldridge (2002, ch. 15.10).

Table 1: Analysis of rank in Google list of results

| | Dependent categorical variable: rank in Google list of results | | | | | |
| | Rank 1-10 | | Rank 11-20 | | Rank $\geq$ 21 | |
	m.eff.	s.e.	m.eff.	s.e.	m.eff.	s.e.
Width of product range $*10^{-2}$	0.079***	0.019	0.124***	0.027	-0.203***	0.040
Relative price level	0.058***	0.022	0.091***	0.027	-0.149***	0.047
Online advertising:						
Number of banner ads	0.010***	0.003	0.015***	0.005	-0.025***	0.007
Number of sponsored links	0.003**	0.001	0.004**	0.002	-0.007**	0.003
Retailer service:						
Convenient navigation	-0.010***	0.004	-0.015***	0.006	0.025***	0.008
Superior customer service	-0.012	0.011	-0.019	0.017	0.032	0.028
Return policy + lens care	-0.010	0.017	-0.016	0.027	0.025	0.043
Security + trustworthiness	-0.019	0.013	-0.030	0.020	0.049	0.032
New CL user service	-0.002	0.005	-0.004	0.008	0.006	0.013
Thresholds:						
μ_1 (s.e.)	3.767 (0.586)					
μ_2 (s.e.)	4.260 (0.568)					
Number of observations	929					
Log likelihood	-257.4					
Wald test (χ_9^2)	143.2					
Pseudo R^2	0.230					

Notes: Marginal effects of ordered probit estimation. The marginal effects are calculated at the mean values of continuous variables. ***, **, * denotes significance at the 1%, 5% and 10% level.

with offering more favourable service. Also cheaper and more specialised e-tailers (offering a narrower product range) are significantly more often found below the first 20 entries in the list of results. E-tailer's investment in online advertising is however negatively correlated with an adverse rank.

The causality of the observed relationship between an e-tailer's rank in the Google list of results and its online advertising activities could, however, go in both directions. On the one hand, e-tailers could use online advertising as a complement to a high search engine rank. On the other hand, it is coceivable that online advertising is used as a means to overcome an unfavourable search engine rank. In which way the causality goes, can not be disentangled using the data at hand. A higher rank being associated with a higher price level as compared to the competitors could stem from a reverse causality in the sense that online retailers with a superior search engine rank are possibly able to exploit this prominent position through price mark-ups. This possibility is suggested by Smith et al. (2000).

5.2.2 Determinants of the Advertising Decision

In order to capture an additional dimension of the concept of virtual location, the use of different types of online advertising is explained by the profit-maximising strategy of e-tailers and the search engine rank which can be expected to be exogenous to the advertising decision. Thus, the estimated effects are interpreted in a causal sense. The fundamental decision with respect to online advertising is analysed, implying that the observed outcome is a categorical variable, indicating the type of online advertising strategy.[13]

The dependent variable distinguishes between online retailers investing only in banner ads, those investing only in sponsored links, online retailers investing in both forms of online advertising, and those without any online advertising.[14] The frequency distribution of the online advertising categories can be found in Table 7 in the Appendix. The advertising decision is explained by the width of the product range, the relative price level, the Google rank and the factors describing the e-tailer services.

On the contrary to the search engine ranks, the categories describing the online advertising decision are not ordered but only mutually exclusive. The latent profit index idea of Equation 1 can be specified leading to the multinomial logit model (see for example Wooldridge, 2002, ch. 15.9). The multinomial logit model is based on the assumption that the observed outcome y is the one which the individual (or firm) attaches the largest latent utility (or here: profit) to. The probability of observing alternative j can in our case with three possible forms of advertising be described as:

$$Pr(y = j \mid \mathbf{x}) = \frac{exp(\mathbf{x}\beta_j)}{1 + \sum_{h=1}^{3} exp(\mathbf{x}\beta_h)}$$

where $j = 1, 2, 3$. Once the probabilities for $j = 1, 2, 3$ are specified, the probability for no online advertising $Pr(y = 0 \mid \mathbf{x})$ is known, because the probabilities must sum to unity. The parameter vector β is estimated by maximum likelihood. The estimated coefficients and the marginal effects $\partial p_j(\mathbf{x})/\partial x_k$ computed separately for each outcome category are denoted together with their estimated standard errors in Table 2.

---

[13]The actual numbers of banner ads or sponsored links are thus neglected.

[14]Note that only key word-related advertising in Internet search engines is observed in the data. Other types of activities such as banner ads on portals or general interest websites, for example, are neglected.

Table 2: Analysis of online advertising activities

| | Dependent categorical variable: type of online advertising activities | | | | | | | |
| | None | | Banner ads | | Sponsored links | | Both | |
	m.eff.	s.e.	m.eff.	s.e.	m.eff.	s.e.	m.eff.	s.e.
Width of product range $*10^{-2}$	-0.019	2.112	0.004	0.906	0.018	1.663	-0.002	0.435
Relative price level	0.270***	0.085	-0.168**	0.075	-0.035	0.025	-0.067**	0.027
Google rank (ref.: ≥ 21):								
Google rank 1-10	-0.231***	0.084	0.093*	0.053	0.026	0.030	0.112*	0.059
Google rank 11-20	-0.329***	0.104	0.042	0.033	0.091	0.065	0.196**	0.094
Retailer service:								
Convenient navigation	-0.030***	0.008	0.003	0.004	0.023***	0.007	0.005**	0.002
Superior customer service	0.009	0.008	-0.015*	0.008	0.005*	0.003	0.000	0.002
Return policy + lens care	-0.038	0.030	0.015	0.013	0.011	0.010	0.012	0.010
Security + trustworthiness	-0.010**	0.005	0.003	0.003	0.003	0.003	0.004*	0.002
New CL user service	0.009	0.034	-0.002	0.014	-0.007	0.013	-0.001	0.009
Number of observations	929							
Log likelihood	-372.29							
Wald test (χ^2_{27})	306.46							
Pseudo R^2	0.3237							

Notes: Marginal effects of multinomial logit estimation. The marginal effects of continuous variables and the base categories of dummies. The marginal effects of dummy variables are calculated as discrete changes in the expected value of the dependent variable. ***, **, * denotes significance at the 1%, 5% and 10% level.

The width of the product range offered by online retailers has no significant influence on their decision to promote their web shop online. By contrast to this, the relative price level of e-tailers significantly influences the online advertising strategy: The more expensive an e-tailer is, the more probably it does not invest in online advertising. Moreover, e-tailers being ranked higher in the Google list of results invest more probably in both types of online advertising, whereas e-tailers with less favourable ranks do not invest in online advertising with a larger probability. This suggests that e-tailers tend to complement a superior Google rank with online advertising, further enhancing their visibility on the web and thus increasing the probability of being found by potential customers.

Being relatively cheaper in comparison to the competitors significantly increases the probability of investing at least in banner ads. This implies that banner ads seem to be used as a means of signalling low prices.[15] On the contrary, the positive marginal effects of the variables describing retailer services which are estimated for the probability of using sponsored links indicate that sponsored links seem to serve as a vehicle for signalling superior customer service. Both results appear plausible, when the different optical designs of banner ads and sponsored links are considered. Banner ads are the more aggressive type of advertising because they are graphically highlighted and easily distinguishable from the list of search results. If relatively cheaper e-tailers use banner ads with a higher probability, this result can be interpreted as evidence for banner ads targeting fairly price-sensitive consumers which respond to this type of advertising design. Besides, sponsored links which are designed to hide on top of the list of search results and which differ only slightly from ordinary search results, can be thought of as appealing to consumers searching for a reliable supplier or for superior e-tailer quality instead of hunting for a bargain.

In order to check the robustness of the results, various further specifications of the multinomial logit model have been estimated. The estimated coefficients of these specifications are denoted in Table 3 together with the coefficients of the original specification which is equivalent to specification (2). The standard errors are estimated by nonparametric bootstrapping in order to account for the inclusion of generated regressors, which were obtained by the factor analysis. The first specification explains the online advertising decision by the same explanatory variables except the 5 factors describing e-tailer services. Specification (3) uses the one-month lag of the search engine rank instead of the current rank to avoid endogeneity of the Google rank. In the fourth specification, an additional variable is included, describing if the e-tailer is a "pure e-tailer" – meaning an e-tailer being active only in the online business and not running an additional physical store. These results are fairly robust to these different specifications, and it must be noted that the estimated effects of the search engine rank are robust to different ways of accounting for e-tailer quality attributes.

---

[15]Banners were only sporadically used for price advertising during the period of data collection.

Table 3: Robustness check

	Dependent categorical variable: type of online advertising activities							
	(1)		(2)		(3)		(4)	
	coeff.	s.e.	coeff.	s.e.	coeff.	s.e.	coeff.	s.e.
Only banner ads								
Width of product range $*10^{-2}$	2.076***	0.639	0.461	0.926	0.202	1.181	0.303	1.185
Relative price level	-13.117***	2.123	-20.368***	4.607	-20.595***	5.517	-20.811***	5.766
Pure e-tailer							0.248	0.371
Google rank (ref.: ≥ 21):								
Google rank$_t$ 1-10	1.960	1.508	2.877*	1.514				
Google rank$_t$ 11-20	1.640	6.138	2.299	6.077				
Google rank$_{t-1}$ 1-10					2.716***	0.762	2.670***	0.789
Google rank$_{t-1}$ 11-20					2.334	8.066	2.301	7.773
Retailer service:								
Convenient navigation			0.376*	0.213	0.399*	0.220	0.402*	0.223
Superior customer service			-1.770***	0.668	-1.773***	0.603	-1.903***	0.662
Return policy + lens care			1.837*	1.099	1.897**	0.830	1.908**	0.869
Security + trustworthiness			0.364	0.238	0.341	0.250	0.349	0.257
New CL user service			-0.195	1.032	-0.338	0.509	-0.303	0.511
Constant	9.468***	1.908	15.724***	3.926	15.946***	4.701	16.053***	4.895
Only sponsored links								
Width of product range $*10^{-2}$	2.476**	1.102	1.400	1.502	1.567	1.588	1.320	1.612
Relative price level	-5.757***	1.358	-2.971	1.946	-2.830	2.011	-2.644	2.315
Pure e-tailer							-0.383	0.767
Google rank (ref.: ≥ 21):								
Google rank$_t$ 1-10	1.426	8.768	1.398	8.870				
Google rank$_t$ 11-20	2.141	1.976	2.585	2.064				
Google rank$_{t-1}$ 1-10					1.657	8.618	1.569	8.788
Google rank$_{t-1}$ 11-20					2.955	1.952	2.911	1.818
Retailer service:								
Convenient navigation			1.777***	0.430	1.846***	0.409	1.868***	0.466
Superior customer service			0.378	0.252	0.327	0.251	0.385	0.319
Return policy + lens care			0.864	0.533	0.915**	0.435	0.915*	0.469
Security + trustworthiness			0.272	0.253	0.323	0.229	0.312	0.292
New CL user service			-0.522	0.620	-0.553**	0.249	-0.603**	0.259
Constant	1.715	1.423	-1.800	2.186	-2.039	2.228	-2.071	2.540
Both								
Width of product range $*10^{-2}$	1.730**	0.814	-0.491	0.934	0.188	1.082	2.356**	1.051
Relative price level	-10.349***	1.487	-14.002***	2.465	-14.829***	3.177	-14.599***	3.696
Pure e-tailer							2.264***	0.607
Google rank (ref.: ≥ 21):								
Google rank$_t$ 1-10	2.771***	0.537	3.600***	0.697				
Google rank$_t$ 11-20	2.933***	0.653	4.314***	0.809				
Google rank$_{t-1}$ 1-10					3.683***	0.849	3.118***	0.849
Google rank$_{t-1}$ 11-20					4.715***	0.931	3.913***	0.902
Retailer service:								
Convenient navigation			1.003**	0.432	1.228***	0.440	1.426***	0.495
Superior customer service			0.043	0.353	0.032	0.411	-0.237	0.321
Return policy + lens care			2.440*	1.371	2.688***	0.869	2.997***	0.827
Security + trustworthiness			0.842**	0.390	1.003***	0.361	0.825**	0.397
New CL user service			-0.258	1.287	-0.371	0.518	-0.154	0.631
Constant	6.260***	1.519	8.722***	2.490	8.970***	3.173	6.952**	3.516
Number of observations	929		929		779		779	
Log likelihood	-457.19		-372.29		-306.11		-296.23	
Wald test (χ^2_{27})	173.78***		306.46***		261.82***		256.70***	
Pseudo R^2	0.1695		0.3237		0.3385		0.3599	

Notes: Estimated coefficients of multinomial logit estimation with "no online advertising" as comparison group. The marginal effects are calculated at the mean values of continuous variables and the base categories of dummies. The marginal effects of dummy variables are calculated as discrete changes in the expected value of the dependent variable. ***, **, * denotes significance at the 1%, 5% and 10% level.

Further checks, such as splitting the sample into a high- and a low-quality subsample and then estimating separately on both samples, could not be performed due to the sample size. The validity of the independence of irrelevant alternatives (IIA) assumption has been checked by a Small/Hsiao test which does not reject the IIA assumption. The results are available from the author on request.

Some effort has been put into the question whether the search engine rank is endogenous with respect to the advertising strategy, that is, whether the Google rank can really be treated as an exogenous variable in Table 2. In order to gain some insight into the problem two test strategies[16] both being based on an instrumental variable approach have been performed. The width of the product range offered by the e-tailers appears to be usable as an instrument for the search engine rank, as it turned out to have no effect on the advertising strategy in Table 2 but to affect the search engine rank even after conditioning on other covariates, as can be seen in Table 1.

A plausible interpretation for this result stems from the unknown mechanism by which Google ranks its results: online shops offering a wider product range tend to consist of a larger bundle of websites than online shops offering less products. This possibly affects the number of links which is used by Google as one criterion among others to rank the search results. Besides, there is no clear theoretical relationship between the number of products offered by an e-tailer and its advertising strategy, as specialized suppliers offering a limited product variety as well as e-tailers offering the full range of available products may both invest in online advertising or not.

As the width of the product range is the only available potential instrumental variable the Google rank is (incorrectly) treated as a continuous variable being censored at 101 (using dummy variables describing the Google rank, one instrumental variable would have been necessary for each dummy variable). First, the exogeneity test by Smith and Blundell (1986) for the probit model is applied with the null hypothesis of exogeneity. For performing this test all advertising strategies are tested separately, which is possible due to the independence of irrelevant alternatives assumption. Second, a simple plug-in test (analogous to a 2SLS approach) is applied where in a first stage the Google rank is regressed on the covariates and the instrumental variable (see Wooldridge, 2002, p. 474). In the second stage, a multinomial logit augmented by the residual from the first stage regression is estimated. The null hypothesis is then tested by a t-test on the coefficient of the residual. Using both test strategies, the null hypothesis of endogeneity can be rejected and the search engine rank can be used as an exogenous variable in the analysis of the determinants of the online advertising strategy. The results are available from the author on request.

5.2.3 Changes in the Virtual Location

In the next step it would be interesting to investigate changes in the virtual location, particularly in the search engine rank. The frequency distribution of changes in the

---

[16]Both tests were originally developed for different limited dependent variable models but are also applicable to the multinomial logit case.

Google rank is described in Table 4. It can be seen that substantial changes in the rank do not occur very often. Most e-tailers stay in the same rank category in which they were ranked in the previous month. Of course, changes occur in the exact list of results, but who would notice if a seller with position 56 in May would climb to position 53 in June? Unfortunately, the number of substantial changes in the Google results is too small to be analysed econometrically. It can only be concluded that search engine ranks appear to remain relatively stable over time. This relatively low turnover in search engine ranks suggests that the market was in equilibrium during the period of observation.

Table 4: Changes in the search engine rank

rank in month t	rank in month $t+1$		
	1-10	11-20	≥ 21
1-10	27	5	2
11-20	7	16	12
≥ 21	1	14	697

6 Conclusions

In this paper, a first empirical analysis of the profit maximising strategies of e-tailers with respect to their virtual location has been provided. The important role of an outstanding virtual location in the attraction of new customers has been discussed.

The empirical results suggest that for e-tailers optimising their strategy with respect to their virtual location, a high Google rank and online advertising appear to be complements, as e-tailers being ranked among the first 20 results decide to use banner ads as well as sponsored links with a higher probability than e-tailers being adversely ranked. Moreover, banner ads seem to serve as price advertising mechanism targeting price-sensitive customer groups, whereas sponsored links rather seem to be used in order to signal outstanding customer service to more quality-oriented consumers. A test strategy using the width of product range as instrumental variable has revealed that the search engine rank can indeed be assumed to be exogenous in the analysis of the online advertising strategy.

To complete the descriptive picture on e-tailers and their virtual location, there must be kept in mind that there is a significant correlation between the combination of relatively high price level/wider product range/inferior customer service, and a high search engine rank. Considering the development over time, the search engine rank appears to remain relatively stable. This suggests that the market was in equilibrium during the period of observation.

Up to now, the literature on the virtual location of firms has been sparse. The role of promotional (or brand) advertising online has been illustrated by Baye and Morgan (2004), and this is the only empirical study examining promotional advertising in the online environment which is known to the author of this paper. Future

research should focus on the competing roles of promotional and informational (or price) advertising: What exactly is the trade-off for e-tailers when allocating their budget for online advertising? In this context – but also for the isolated analysis of virtual location – information on consumer response to online advertising would be quite useful. This lack of information could be alleviated using clickstream data (like for example in Baye, Gatti, Kattuman and Morgan (2004), Smith and Brynjolfsson (2001) or Goldfarb (2002)). Using such data would alleviate the problem but probably not solve it completely, as only click-throughs to an e-tailer's website can be observed, but not actual purchasing decisions or consumer characteristics.

References

Baye, M. R., Gatti, J. R. J., Kattuman, P. and Morgan, J. (2004). Estimating Firm-Level Demand at a Price Comparison Site: Accounting for Shoppers and the Number of Competitors, *mimeo*, download: http://www.econ.cam.ac.uk/faculty/gatti/wip.html.

Baye, M. R. and Morgan, J. (2001). Information Gatekeepers on the Internet and the Competitiveness of Homogeneous Product Markets, *American Economic Review* **91**(3): 454–474.

Baye, M. R. and Morgan, J. (2004). Brand and Price Advertising in Online Markets, *mimeo*, download: http://www.nash-equilibrium.com/baye/Promotion.pdf.

Baye, M. R., Morgan, J. and Scholten, P. (2004). Price Dispersion in the Small and in the Large: Evidence from an Internet Price Comparison Site, *Journal of Industrial Economics* **52**(4): 463–496.

Benham, L. (1972). The Effect of Advertising on the Price of Eyeglasses, *Journal of Law & Economics* **15**(2): 337–352.

Bhatnagar, A. and Papatla, P. (2001). Identifying Locations for Targeted Advertising on the Internet, *International Journal of Electronic Commerce* **5**(3): 23–44.

Chatterjee, P., Hoffman, D. L. and Novak, T. P. (2003). Modeling the Clickstream: Implications for Web-Based Advertising Efforts, *Marketing Science* **22**(4): 520–541.

Clay, K., Krishnan, R., Wolff, E. and Fernandes, D. (2002). Retail Strategies on the Web: Price and Non-Price Competition in the Online Book Industry, *Journal of Industrial Economics* **50**(3): 351–367.

ECIN (2003a). Online-Werbung in der Krise? Download: www.ecin.de/news/2003/11/25/06466/.

ECIN (2003b). Suchmaschinen-Marketing: ROI in Echtzeit. Download: www.ecin.de/marketing/suchmaschinenreoi/.

Ellison, G. and Fisher Ellison, S. (2004). Search, Obfuscation, and Price Elasticities on the Internet, *Discussion Paper 10570*, NBER, Cambridge, Massachusetts.

Fittkau & Maaß (2003). *17. Nutzeranalyse W3B*, Fittkau & MaaßConsulting GmbH, Hamburg.

Goldfarb, A. (2002). Analyzing Website Choice Using Clickstream Data, *in* M. R. Baye (ed.), *The Economics of the Internet and E-Commerce*, number 11 in *Advances in Applied Microeconomics*, JAI.

Häring, J. (2003). Different Prices for Identical Products? Market Efficiency and the Virtual Location in B2C E-Commerce, *Discussion Paper 03-68*, Centre for European Economic Research, Mannheim.

Hoffman, D. L. and Novak, T. P. (2000). Advertising Pricing Models for the World Wide Web, *in* B. Kahin and H. R. Varian (eds), *Internet Publishing and Beyond: The Economics of Digital Information and Intellectual Property*, MIT Press, Cambridge, Massachusetts, pp. 45–61.

Hunter, D. (2003). Cyberspace as Place and the Tragedy of the Digital Anticommons, *California Law Review* **91**(2): 439–519.

Kenny, D. and Marshall, J. F. (2000). Contextual Marketing: The Real Business of the Internet, *Harvard Business Review* **78**(6): 119–125.

Kwoka, J. E. (1984). Advertising and the Price and Quality of Optometric Services, *American Economic Review* **74**(1): 211–216.

Luo, X. and Seyedian, M. (2004). Contextual Marketing and Customer-Orientation Strategy for E-Commerce: An Empirical Analysis, *International Journal of Electronic Commerce* **8**(2): 95–118.

Machill, M., Neuberger, C., Schweiger, W. and Wirth, W. (2003). Wegweiser im Netz: Qualität und Nutzung von Suchmaschinen, *in* M. Machill and C. Welp (eds), *Wegweiser im Netz: Qualität und Nutzung von Suchmaschinen*, Verlag Bertelsmann Stiftung, Bielefeld, pp. 13–490.

Manchanda, P., Dubé, J.-P., Goh, K. Y. and Chintagunta, P. K. (2004). The Effect of Banner Advertising on Internet Purchasing, *mimeo*, University of Chicago, Graduate School of Business, Chicago.

McFadden, D. (1974). Conditional Logit Analysis of Qualitative Choice Behavior, *in* P. Zarembka (ed.), *Frontiers of Econometrics*, Academic Press, New York.

Morgan, J., Orzen, H. and Sefton, M. (2003). A Laboratory Study of Advertising and Price Competition, *Technical Report 2003-04*, University of Nottingham, Centre for Decision Research and Experimental Economics, Nottingham.

Silk, A. J., Klein, L. R. and Berndt, E. R. (2001). The Emerging Position of the Internet as an Advertising Medium, *Netnomics* **3**(2): 129–148.

Smith, M. D. (2002). The Law of One Price? The Impact of IT-enabled Markets on Consumer Search and Retailer Pricing, *Working Paper 2002-35*, H. John Heinz III School of Public Policy and Management, Carnegie Mellon University, Pittsburgh. http://www.heinz.cmu.edu/wpapers/retrievePDF?id=2002-35.

Smith, M. D., Bailey, J. and Brynjolfsson, E. (2000). Understanding Digital Markets: Review and Assessment, *in* E. Brynjolfsson and B. Kahin (eds), *Understanding the Digital Economy*, MIT Press, Cambridge, Massachusetts, pp. 99–136.

Smith, M. D. and Brynjolfsson, E. (2001). Consumer Decision-Making at an Internet Shopbot: Brand Still Matters, *Journal of Industrial Economics* **49**(4): 541–558.

Smith, R. J. and Blundell, R. W. (1986). An Exogeneity Test for a Simultaneous Equation Tobit Model with an Application to Labor Supply, *Econometrica* **54**(3): 679–686.

Stahl, D. O. (2000). Strategic Advertising and Pricing in E-Commerce, *in* M. R. Baye (ed.), *Industrial Organization*, Vol. 9 of *Advances in Applied Microeconomics*, JAI, Amsterdam, pp. 69–100.

Tang, F.-F. and Lu, D. (2001). Pricing Patterns in the Online CD Market: An Empirical Study, *Electronic Markets* **11**(3): 171–185.

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*, MIT Press, Cambridge, MA.

Appendix

Figure 2: Banner ads and sponsored links in a Google list of results
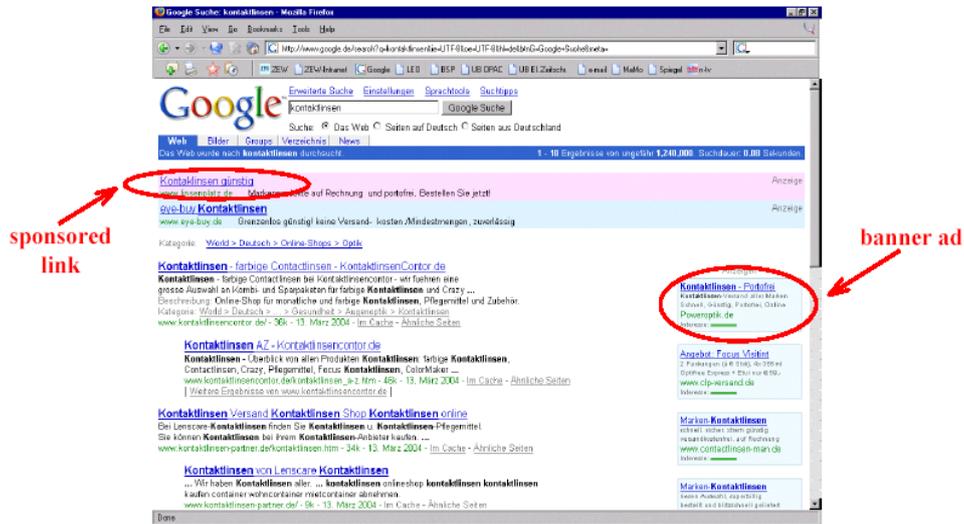
Note: Screenshot from 14.03.2004

Figure 3: Banner ads and sponsored links in a Lycos list of results
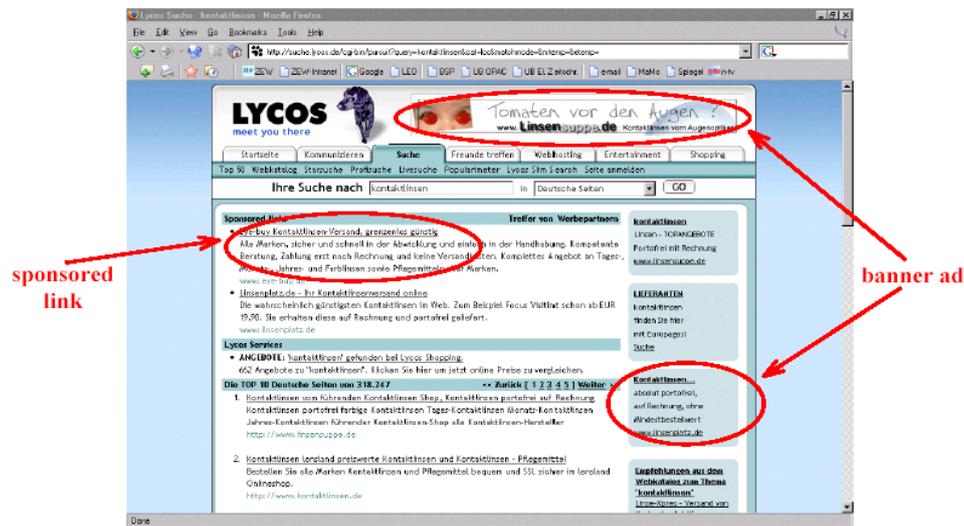
Note: Screenshot from 14.03.2004

Table 5: 10 most-widely used search engines

www.google.de
www.yahoo.de
search.msn.de
www.lycos.de
www.t-online.de
www.altavista.de
www.web.de
www.metager.de (meta search engine)
www.fireball.de
suche.aol.com

Source: www.webhits.de, 04.03.2002.

Table 6: Frequency distribution of Google rank categories

Google rank	Frequency
Google rank 1-10	41
Google rank 11-20	41
Google rank $\geq$ 21	847
Number of observations	929

Table 7: Frequency distribution of online advertising categories

Online advertising activities	Frequency
Banner ad	66
Sponsored link	34
Both	43
None	786
Number of observations	929

Table 8: Descriptive Statistics of explanatory variables

	Mean	Std. Dev.
Width of product range	21.743	17.661
Relative price level	1.024	0.135
Number of banner ads	0.300	1.062
Number of sponsored links	0.670	2.917
Google rank (ref.: not ranked 1-100):		
Google rank 1-10	0.044	
Google rank 11-20	0.044	
Google rank ≥ 21	0.912	
Retailer service:		
Convenient navigation	0.015	0.844
Superior customer service	0.002	0.813
Return policy + lens care	0.018	0.810
Security + trustworthiness	0.008	0.813
New CL user service	0.011	0.794
Number of observations	929	