

Fabel, Oliver; Lehmann, Erik E.

**Working Paper**

## Adverse selection and the economic limits of market substitution: An application to e-commerce and traditional trade in used cars

Diskussionsbeiträge - Serie I, No. 302

**Provided in Cooperation with:**

Department of Economics, University of Konstanz

*Suggested Citation:* Fabel, Oliver; Lehmann, Erik E. (2000) : Adverse selection and the economic limits of market substitution: An application to e-commerce and traditional trade in used cars, Diskussionsbeiträge - Serie I, No. 302, Universität Konstanz, Fachbereich Wirtschaftswissenschaften, Konstanz

This Version is available at:

<https://hdl.handle.net/10419/68880>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

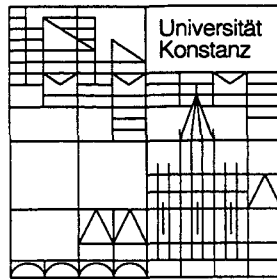
Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

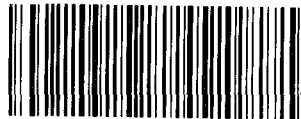


402 972 203

**Rechts-, Wirtschafts- und  
Verwaltungswissenschaftliche Sektion**  
Fachbereich  
Wirtschaftswissenschaften

---

W 284 (302)



Oliver Fabel  
Erik E. Lehmann

**Adverse Selection  
and the Economic Limits  
of Market Substitution:  
An Application to E-Commerce  
and Traditional Trade  
in Used Cars**

24. MAZ. 2000 Weltwirtschaft  
Kiel

W 284 (302)

Du es sig. 97.5 Diskussionsbeiträge

**Adverse Selection and the Economic  
Limits of Market Substitution:  
An Application to E-Commerce and Traditional  
Trade in Used Cars**

**Oliver Fabel and Erik E. Lehmann**

Serie I – Nr. 302

**Februar 2000**

### **Abstract:**

Adverse selection induces economic limits to market substitution. If quality uncertainty persists in both internet and traditional marketplaces, a second-best equilibrium with parallel market segments may arise. Positive trade in parallel segments implies that the information cost advantage of one marketplace is exactly offset by a more severe adverse selection problem associated with non-observable quality variables. The electronic marketplace providing dominant search means contains all segments, while the traditional market may lack some segments. These missing segments are characterized by low quality expectations given the vector of advertised quality signals. The analytic results are confirmed by an empirical investigation of used-car trade. Thus, the study also provides an estimate of the price differential between the electronic and the traditional marketplace.

**Keywords:** E-Commerce, Market Substitution, Adverse Selection, Efficient Search and Learning.

**JEL-Classification:** Business Economics (M21), Market Structure, Firm Strategy and Market Performance (L1), Asymmetric and Private Information (D82), Search, Learning and Information (D83).

Authors' address:

Prof. Dr. Oliver Fabel and Dr. Erik Lehmann, University of Konstanz, Faculty of Law, Economics, and Public Administration, Department of Economics, Chair for Management, in particular Firm Policy Formation, Box D144, 78457 Konstanz, Germany.

Phone: +49-(0)7531-88-2990/2992 (Fabel), +49-(0)7531-88-2968 (Lehmann).

Fax: +49-(0)7531-88-4456.

E-Mail: [oliver.fabel@uni-konstanz.de](mailto:oliver.fabel@uni-konstanz.de), [erik.lehmann@uni-konstanz.de](mailto:erik.lehmann@uni-konstanz.de).

# 1 Introduction

Following Barua, et al. (1999), OECD (1999a), and Mesenbourg (1999), defining and measuring the economically relevant impacts of internet activities remains an open issue. In particular, this is seen to apply to the development of "e-commerce". There exist various measures of business-to-business and business-to-consumer trade in common goods and services. However, according to all recent policy reports - such as US Department of Commerce (1999), OECD (1997; 1999b, ch. 1), and the Gemini (1999) study initiated by the European Commission, for instance - the respective social product contributions must currently be viewed as insignificant. Hence, estimates of the percentage of internet mediated sales in total retailing value range between less than 1 % [US Department of Commerce (1999)] and less than 0.5 % [OECD (1999b, p. 27) for seven countries]. Moreover, Gemini (1999, p.1) expects the impact of such internet activities on the European economies' domestic products and labor markets to remain "*relatively small*" for the next five years.

Policy reports therefore rather emphasize the extreme growth rates realized in this sector. Furthermore, the studies generally agree in predicting significant spill-over effects fostering general economic growth. OECD (1999b) thus expects the value of e-commerce activities to grow from \$ 26 billion in 1997 to \$ 330 billion in 2001-02 and \$ 1 trillion in 2003-05. Most recent figures by Internet Economy Indicators (1999) - a detailed description of which can again be found in Barua, et al (1999) - estimate that commercial internet revenues experienced 127 % growth between 1998 and 1999. They already account for 26 % of the \$ 116.832 billion total revenue generated in the so-called "internet economy". The respective 78 % employment growth in e-commerce also well exceeds the 46 % industry average. In 1999, 900,882 employees in commercial internet firms constituted 39 % of total internet related employment. This compares to 28 % employed in structuring and maintaining the technical infrastructure, 24 % in internet software design and consultancy, and 19 % in the provision of intermediation services such as search engines, electronic brokerage, and auctioning. Thus, e-commerce will "*expand existing markets and even create new ones*" [US Department of Commerce (1999, p. 10)]. At the same time, it "*can significantly influence existing market activities*" [OECD (1997, p. 44)]. In particular, the "*the longer-term viability of firms [...] will be determined by the strategies they choose to pursue*" [Gemini (1999, p. 1)].

However, the internet economy has only begun to attract the attention of the economics discipline. Also, the existing literature almost exclusively focuses on the specific aspects of internet technologies, services, and products. The contributions collected in Kahin and Keller (1995) and McKnight and Baily (1997) thus discuss the volume and structure of internet communication, problems of congestion pricing, tariffs for internet services, and the distribution of internet access over the population. Even marketing studies - unambiguously market and strategy oriented - are currently entirely confined to investigating the impacts of internet technology on consumer and supplier behavior. The pace and unpredictability of technological change then either precludes the possibility of projecting future developments as in Peterson, et al. (1997), or inspires conceptual ideas concerning the design of the future net environment in Hoffman and Novak (1996). Hence, despite the obvious business and public interest, analyses of the interdependencies between traditional and new internet markets are generally lacking.

Clearly, internet technology and services alter the search behavior of potential buyers. Thus, Hoque and Lohse (1999) demonstrate that successful consumer search implies different strategies and costs in traditional and internet "yellow pages". Moreover, policy reports [US Department of Commerce (1999, p. 10-12), Gemini (1999, p.13, 17-18), OECD (1997, p. 6; 1999b, ch. 2)], economic analysis [Zerdick, et al. (1999, p. 146-154), Choi, et al. (1997, Ch. 7)], and marketing studies [Alba, et al.(1997)] agree that internet technology reduces search costs. Suppliers, internet operators, and specialized service firms invest significantly in the provision of efficient search means. Software programs which operate as agents of the consumers scan all available information concerning prices and quality attributes of internet offers. Such "shopbots" - analyzed by Kephart and Greenwald (1999), Greenwald and Kephart (1999), and Greenwald, et al. (1999) - increase the price elasticity of demand. Further, in addition to the "globalization" effect, competition among firms is also enhanced by new "non-industry" entrants [OECD (1999b, p. 80-82)]. Consequently, internet trade should increase consumer surplus [Bakos (1997)].

This obviously raises the question whether traditional markets may remain viable at all. Given the technological development, respective approaches emphasizing the transaction costs of entering internet and traditional markets appear myopic. However, the existing internet economics literature also provides some hints concerning the possible longer-term limits of market substitution. Hence, Greenwald and Kephart (1999) and Greenwald, et al. (1999) already investigate the strategic interaction of

"shopbot" consumer and "pricebot" supplier agents in internet markets. The latter employ price-setting algorithms responding to the "shopbot" queries. Further, regulators worldwide - as summarized in US Department of Commerce (1999), Gemini (1999) and OECD (1997) - currently attempt to achieve a balance between supporting the development of the internet economy and consumer protection. They express significant concerns associated with the security and privacy of e-commerce transactions, as well as domain name identification and quality standards for sellers. Finally, Choi, et al. (1997, ch. 8) and Kalakota and Whinston (1999) emphasize that successful internet marketing necessarily requires increased differentiation of suppliers and products. In particular, firms must pick an appropriate mix of market segments.

In summary, these arguments suggest that e-commerce provides dominant search means as far as standardized product and price information is concerned [Bakos (1997)]. However, given the sellers' or product characteristics such standardization may not be possible. Zerdick, et al. (1999, p. 159-160) and Choi, et al. (1997, ch. 5) point out that the usual adverse selection problem is reinforced by network economies. Further, the information provided by suppliers pursuing a differentiation strategy - as illustrated by Choi, et al. (1997, ch. 6) - obviously rather deters the possibilities of comparing offers. Hence, some information cannot, or will not be communicated verifiably. The analysis of market substitution and interaction must therefore explicitly distinguish product or agent information which can be gathered more efficiently via the internet from quality signals which remain private information.

Choi, et al. (1997, p. 169) therefore conclude that the future of the internet "*may depend on how non-technological but fundamentally economic issues as the lemons problem are solved*". Bakos' (1997) analysis of electronic marketplaces already demonstrates that lower search costs may restore, or improve equilibrium trade which is subject to quality uncertainty. However, the current model specifically addresses the issue of market substitution. It demonstrates that, if adverse selection cannot be eliminated completely, internet and traditional markets may co-exist. The analysis draws on standard adverse selection theory. However, it also introduces a risk reducing, but not eliminating linear learning process. The costs of gathering informative quality signals then vary between marketplaces. Nevertheless, an equilibrium with parallel markets or market segments may exist, if the migration of sellers reduces the adverse selection problem in the market characterized by higher information costs. While the migration process between markets - hence, the evolutionary dynamics - cannot be modelled, the

equilibrium is robust with regard to both seller and buyer incentives. The analysis thus also contrasts with approaches exclusively emphasizing the dynamic aspects of internet market emergence.

Obviously, the formal model combines well-known pieces of economic theory. Nevertheless, the empirical application to trade in used-cars pursued in section 3 does not only reflect an interest in testing Akerlof's (1970) seminal work on the "lemon market". The quality of a used car is determined by investments of both original manufacturers and current owners in technical reliability. However, the maintenance effort of the latter remains private information. In addition, the typical spot-market characteristics preclude formal warranty or "money back" agreements as discussed in Choi, et al. (1997, Ch. 4). Hence, in many respects the used-car market - although predominantly involving consumer-to-consumer-trade - provides a prime example for investigating the impact of private seller information on retailing in experience goods. E-commerce activities then only provide an alternative means of exchange [Bakos (1997), Keeney (1999)]. Furthermore, internet trade is already well-developed in both the classified ads and automotive businesses. According to the US Department of Commerce (1999, p. 10), Gemini (1999, p. 21), and OECD (1999b, p. 44, 48), this implies a significant degree of market substitution. As a by-product, classified ads for used cars therefore allow to compile data for empirical analysis.

## 2 A model framework

### 2.1 Quality and information structure

Given the empirical application to follow, the analytical discussion will be framed within the terms of the used car market. The quality of a used car reflects the value of transportation services minus repair and maintenance costs incurred by its future owner. Formally, let  $q_\ell$  denote the quality of a car in the possession of current owner  $\ell$ . Then,

$$q_\ell = q(\bar{\alpha}, \bar{y}) + \beta_\ell + \varepsilon \quad (1)$$

where  $q(\bar{\alpha}, \bar{y})$  captures the quality effects which are associated with car-type and  $\beta_\ell$  reflects the maintenance effort and ability of its current owner.

More precisely,  $\bar{\alpha}$  constitutes a deterministic vector of car characteristics. This list includes brand name and model type, horse-power, size of engine, but - where



applicable - also the accessory units such as air-conditioning, sun-roof, etc. The vector  $\bar{y}$  then contains signals concerning the car's current state of attrition which are related to the original manufacturer's investments in protection against the average wear and tear. Given the present application, age, mileage, and mileage per year since first licensing constitute typical  $\bar{y}$ -entries - the latter easily calculated from the first two entries. While high age and mileage are certainly "bad news" as far as expectations about future reliability and repair costs are concerned, the third is usually considered to be positively correlated with  $q_\ell$ . High realizations of mileage per year indicate less time spent in the stop-and-go-traffic of cities and more freeway driving.

There further exists an owner-determined quality effect associated with the individual decisions to invest in car maintenance. In accordance with Akerlof (1970),  $\beta_\ell$  constitutes perfectly private information of the current owner. In fact, this variable captures all owner-determined quality enhancing activities which cannot be communicated verifiably to potential buyers. Also,  $\beta_\ell$  is assumed to be independent of  $(\bar{\alpha}, \bar{y})$ . Thus, potential buyers believe that owner-determined quality constitutes the realization of a random variable  $\beta$  which is distributed according to  $H(\beta)$  over the population of current owners. Following Akerlof's (1970) original study and common textbook examples - i.e., Milgrom and Roberts (1992, p. 149-155) and Campbell (1995, p. 184-186) - on adverse selection models,  $H(\beta)$  is taken to be uniform with  $\beta \in [0, \beta^H]$ .

Finally,  $\varepsilon$  denotes a random variable reflecting the stochastic nature of the impacts of both the original manufacturer's and the current owner's investments on realized car quality. For simplicity  $\varepsilon \sim N(0, \sigma^2)$ . The information concerning  $\bar{\alpha}$  and  $\bar{y}$  is communicated via the advertisement placed by owners. In this respect, the used car market found in the classified ad sections of newspapers, but also in the internet has established a market standard. All elements of  $\bar{\alpha}$  and  $\bar{y}$  can be easily verified on eye-sight inspection by a potential buyer. Thus, the analysis excludes the possibility of fraudulent offers associated with deliberate manipulations of the speedometer or title documents by the current owner.

However, there also exists a second set of signals concerning  $q_\ell$ . Hence, motorist clubs and automobile magazines perform endurance and comparative driving tests. They may also provide information concerning typical repairs and statistical information on repair frequency and average repair costs. Similar information may be available from the original manufacturer, or still other sources. In the following, let  $I_\ell$  denote the cardinality of the vector  $\bar{z}_\ell^o$  containing such additional signals concerning  $q_\ell$  which are

known to the current owner. For simplicity,  $\bar{z}_\ell^o$  is assumed to constitute a "complete" signal vector. Current owners have collected all test and other available information concerning the car which has been in their possession over the past years. This implies  $\bar{z}_\ell^o = \bar{z}^o$  for all owners  $\ell$ . It then further appears plausible that  $I = I(\bar{\alpha}, \bar{y})$ . Obviously, more test, endurance, and repair statistics are collected and published as the car ages. Also, certain brands may receive more public attention than others.

Every element  $z_{i,\ell}$ ,  $i = 1, \dots, I(\bar{\alpha}, \bar{y})$ , is taken to satisfy

$$z_{i,\ell} = q_\ell + \eta_i \quad (2)$$

where  $\eta_i$  is normally distributed with expected value zero and identical variance  $\zeta^2$ , for all  $i = 1, \dots, I(\bar{\alpha}, \bar{y})$ . Furthermore,  $cov(\eta_j, \eta_i) = 0$ , for all  $j \neq i$ , and  $cov(\varepsilon, \eta_i) = 0$ , for all  $i = 1, \dots, I(\bar{\alpha}, \bar{y})$ .

Kirby (1993) has already introduced linear Bayesian learning in order to model market processes of merging and separating information pools consisting of signal vectors. Within the current framework this approach allows to distinguish potential buyers' and sellers' expectations concerning product quality. Hence, following De Groot (1970),  $q_\ell$  is then distributed normally with expected value

$$E[q_\ell | \bar{z}^o] = q(\bar{\alpha}, \bar{y}) + \beta_\ell \quad (3)$$

and variance

$$Var[q_\ell | \bar{z}^o] = \left[ \frac{1}{\sigma^2} + \frac{I(\bar{\alpha}, \bar{y})}{\zeta^2} \right]^{-1} \quad (4)$$

from the point of view of a current owner.

In contrast, potential buyers must engage in costly information gathering activities in order to obtain a signal vector  $\bar{z}_m^b(\bar{\alpha}, \bar{y})$  of cardinality  $J_m(\bar{\alpha}, \bar{y}) = |\bar{z}_m^b(\bar{\alpha}, \bar{y})|$ . The information costs are assumed to be given  $C_m(\bar{\alpha}, \bar{y}) = c_m J_m(\bar{\alpha}, \bar{y})$ . Subscript  $m$ ,  $m = \{N, M\}$ , identifies the specific market in which these research activities take place - the internet or the traditional print media classified ad market, respectively. Signals are thus generally perceived as single information units which can be gathered at constant unit costs. Potential buyers must either actively engage in market research or receive the information via communication with current owners prior to purchasing the car. The unit costs of information gathering assumed to be independent depend of whether signals are obtained by own research or communication. Yet, the two market

places differ with respect to the distribution of these costs between the two parties. Thus,  $\gamma_m$ ,  $m = \{N, M\}$ , denotes the share of total information costs incurred directly by the potential buyers. Intuitively,  $\gamma_m C_m(\bar{\alpha}, \bar{y})$  then reflects the costs of market research by potential buyers, while  $(1 - \gamma_m)C_m(\bar{\alpha}, \bar{y})$  can be associated with the costs of communicating signals which are incurred by current owners. This interpretation appears particularly plausible, since current owners must ensure that information is communicated verifiably to potential buyers.

According to Bakos (1997), lower search costs in the electronic marketplaces are mainly associated with gathering price information. The argument is not extended to non-standardized products, or seller information, however. Thus, consider a potential buyer responding to ads in the traditional print media market. Researching on its own the individual will have to scan all relevant publications by motor clubs, manufacturers, etc. Very likely, he will also have to purchase single issues of magazines or test reports. On the other hand, internet shopping implies that - given an initial fixed fee and variable costs which are only related to the duration of the research activities - additional signals can be obtained using the net itself. Providers also supply search machines or "shopbots" for assistance. Yet, the verification costs associated with electronic communication are commonly regarded as relatively high. The internet provides a larger signal pool at the expense of allowing offers into the market without exercising significant filtering. These arguments justify the following assumptions: First,  $c_N < c_M$  captures the general information cost advantage of internet shopping. At the same time,  $\gamma_N > \gamma_M$  reflects the less severe verification problem in communicating information associated with the traditional print media market environment. Thus, potential buyers in the internet engage in more market research on their own.

Clearly, rational buyers engaging in market research will ensure not to collect signals twice. Merely for illustrative reasons suppose that current owners were able to communicate their type  $\beta_\ell$  as well. Then, if a potential buyer in market  $m$  has received the signal vector  $\bar{z}_m^b(\bar{\alpha}, \bar{y})$  consisting of  $J_m(\bar{\alpha}, \bar{y})$  signals,  $E[q_\ell | \bar{z}_m^b(\bar{\alpha}, \bar{y})] = q(\bar{\alpha}, \bar{y}) + \beta_\ell = E[q_\ell | \bar{z}^0]$ . However, the variance equals

$$\text{Var}[q_\ell | \bar{z}_m^b(\bar{\alpha}, \bar{y})] = \left[ \frac{1}{\sigma^2} + \frac{J_m(\bar{\alpha}, \bar{y})}{\zeta^2} \right]^{-1}, \quad m = \{N, M\} \quad (5)$$

Obviously,  $J_m(\bar{\alpha}, \bar{y}) < I(\bar{\alpha}, \bar{y})$  therefore implies that potential buyers are less informed about the technical quality than current owners - even ignoring the owners' private information concerning their investments in car maintenance. The possible

information gap defined by (4) and (5) thus induces an additional risk differential concerning technical reliability. Finally, despite the fact that potential buyers actually collect discrete signals, (5) is taken to be differentiable in  $J_m$  throughout the following analysis.

Note that - as far as the respective stochastic impacts on a used car's quality are concerned - the original manufacturer's and the current owner's investments in reliability are simply additive. Yet, potential buyers actually face two distinct sources of informational deficits relative to current owners. First, different sets of additional informative signals induce differences in the precision of forecasting the car's future value. Potential buyers are able to reduce this information gap by gathering costly, but informative signals, however. The second informational asymmetry arises due to the fact that owner-type information cannot be communicated verifiably at all.

## 2.2 Expected utility and information gathering

The typical classified ad placed by current owners in the used car market contains the general car description  $(\bar{\alpha}, \bar{y})$  and a price for which the owner is willing to sell. Although adjusting to equilibrium different  $(\bar{\alpha}, \bar{y})$ -car owners may advertise various prices, the analysis can be more easily followed by considering only equilibrium ads  $A_m(\bar{\alpha}, \bar{y}) = [P_m(\bar{\alpha}, \bar{y}), \bar{\alpha}, \bar{y}]$  in market  $m$ . In the following, the advertised characteristics  $(\bar{\alpha}, \bar{y})$  will then be said to define segments of market  $m$ . The empirical analysis cannot distinguish between prices contained in advertisements and the terms of actual exchanges. Consistency thus requires that the theoretical model to be developed entails trade triggered by the prices quoted in the advertisements.

The instantaneous preferences of all individuals meeting in both markets will be assumed to be characterized by identical exponential utility functions with constant degree of absolute risk aversion  $r > 0$ . Hence, focussing on the  $(\bar{\alpha}, \bar{y})$ -segment of market  $m$ , an owner of type  $\beta_\ell$  will advertise his car, if

$$-V^s(\bar{\alpha}, \bar{y}) \exp\{-r\beta_\ell\} \leq -\exp\{-r[P_m(\bar{\alpha}, \bar{y}) - (1 - \gamma_m)C_m(\bar{\alpha}, \bar{y})]\} \quad (6)$$

where - exploiting the assumptions above -

$$V^s(\bar{\alpha}, \bar{y}) \equiv \exp\left\{-r \left[ q(\bar{\alpha}, \bar{y}) - \frac{1}{2}r \left( \frac{1}{\sigma^2} + \frac{I(\bar{\alpha}, \bar{y})}{\zeta^2} \right)^{-1} \right] \right\} \quad (7)$$

The LHS of (6) contains the expected utility of retaining the car, while the certain utility derived from the net reward of a sale is given on the RHS. Obviously, (6) - (7) determines a maximum owner-determined car quality offered, given the equilibrium price and the cardinality of the information vector communicated in the  $(\bar{\alpha}, \bar{y})$ -segment of market  $m$ . In particular, only owners characterized by  $\beta_\ell \leq B_m(\bar{\alpha}, \bar{y})$ , where

$$B_m(\bar{\alpha}, \bar{y}) = P_m(\bar{\alpha}, \bar{y}) - (1 - \gamma_m)C_m(\bar{\alpha}, \bar{y}) - q(\bar{\alpha}, \bar{y}) \quad (8)$$

$$+ \frac{1}{2}r \left( \frac{1}{\sigma^2} + \frac{I(\bar{\alpha}, \bar{y})}{\zeta^2} \right)^{-1} \quad (9)$$

are willing to offer their cars for sale.

Potential buyers must incur the cost of market research. Thus, if purchasing a car in the  $(\bar{\alpha}, \bar{y})$ -segment of market  $m$ , a buyers's expected utility is given by

$$U_m(\bar{\alpha}, \bar{y}) = -V_m^b(\bar{\alpha}, \bar{y}) \int_{b_m(\bar{\alpha}, \bar{y})}^{B_m(\bar{\alpha}, \bar{y})} \exp\{-r\beta\} [B_m(\bar{\alpha}, \bar{y}) - b_m(\bar{\alpha}, \bar{y})]^{-1} d\beta \quad (10)$$

where

$$V_m^b(\bar{\alpha}, \bar{y}) \equiv \exp\{-r[\theta + q(\bar{\alpha}, \bar{y}) - P_m(\bar{\alpha}, \bar{y}) - \gamma_m C_m(\bar{\alpha}, \bar{y}) - \frac{1}{2}r \left( \frac{1}{\sigma^2} + \frac{J_m(\bar{\alpha}, \bar{y})}{\zeta^2} \right)^{-1}]\} \quad (11)$$

As in Akerlof's (1970) original work,  $\theta > 0$  induces the possibility of Pareto-improving exchange upon every random match of current owners and potential buyers, if both parties possessed identical information. It reflects the higher urgency of need associated with the non-owner status of an individual. Else, (10) differs from the conventional adverse selection model only in allowing for the possibility that some "poor quality"-owners - characterised by  $\beta_\ell < b_m(\bar{\alpha}, \bar{y})$  - may be absent in this particular segment of market  $m$ .

Focussing on (11), efficient information gathering in market  $m$  obviously implies

$$c_m = \frac{r}{2\zeta^2} \left[ \frac{1}{\sigma^2} + \frac{J_m^*(\bar{\alpha}, \bar{y})}{\zeta^2} \right]^{-2} \quad (12)$$

where  $J_m^*(\bar{\alpha}, \bar{y})$  denotes the optimal activity level. In the following, it will be assumed that (12) implies an interior solution  $J_m^*(\bar{\alpha}, \bar{y}) = J_m^* < I(\bar{\alpha}, \bar{y})$  for all segments

$(\bar{\alpha}, \bar{y})$  in both market environments  $m = \{N, M\}$ . Furthermore, irrespective of the particular means of information gathering - e.g., whether more signals are communicated or obtained by market research of potential buyers - the respective activity levels are always efficient relative to the particular market costs. This solution would obviously always be implemented by optimal market research decisions of potential buyers, if the number of signals actually communicated by current owners in each market does not exceed  $J_m^*$ .

Thus, let  $U_m^*(\bar{\alpha}, \bar{y})$  denote the expected utility of a buyer in segment  $(\bar{\alpha}, \bar{y})$  of market  $m$  given the efficient information gathering activity levels. The alternative for potential buyers always consists of not purchasing a car at all. Hence, if an  $(\bar{\alpha}, \bar{y})$ -segment attracts potential buyers in market  $m$ ,

$$U_m^*(\bar{\alpha}, \bar{y}) \geq -1 \quad (13)$$

must constitute a necessary condition.

### 2.3 Equilibrium with parallel market segments

Consider the internet market  $N$  first. If only this market exists, it is clear that  $b_N(\bar{\alpha}, \bar{y}) = 0$ . Given no possibility to sell in a different market, the gains of entering the used car market decrease with  $\beta_\ell$  - "low quality" owners benefit most from adverse selection. Provided that the segment  $(\bar{\alpha}, \bar{y})$  attracts potential buyers, (8) can then be rearranged to insert for  $P_N(\bar{\alpha}, \bar{y})$  into (11). This yields

$$U_N^*(\bar{\alpha}, \bar{y}) = -V_N^*(\bar{\alpha}, \bar{y}) \int_0^{B_N(\bar{\alpha}, \bar{y})} \exp\{-r\beta\} [B_N(\bar{\alpha}, \bar{y})]^{-1} d\beta \quad (14)$$

with

$$V_N^*(\bar{\alpha}, \bar{y}) = \exp\left\{-r\left[\theta - B_N(\bar{\alpha}, \bar{y}) - c_N J_N^* + \frac{1}{2}r \left( \left[ \frac{1}{\sigma^2} + \frac{I(\bar{\alpha}, \bar{y})}{\zeta^2} \right]^{-1} - \left[ \frac{1}{\sigma^2} + \frac{J_N^*}{\zeta^2} \right]^{-1} \right)\right\} \quad (15)$$

The second line of (15) should be observed to contain the risk differential associated with the information gap between current owners and potential buyers. It is easily verified that (14) is monotonically decreasing in  $B_N(\bar{\alpha}, \bar{y})$ . The attracting of higher quality owners must be achieved by offering higher prices, which will always be attractive for low quality owners as well. Consequently, the adverse selection problem

for buyers will become more severe when the top quality offered in the market increases. Clearly,  $B_N(\bar{\alpha}, \bar{y}) > 0$  also requires that  $\theta$  is sufficiently large in order to compensate the information gap between owners and buyers with regard to technical uncertainty.

Suppose that the internet market involves active trade in more than one  $(\bar{\alpha}, \bar{y})$ -segment. In competitive equilibrium, (13) must then hold with equality for all such segments. This defines equilibrium values  $\tilde{B}_N(\bar{\alpha}, \bar{y})$  for all segments which entail positive trade. However, differentiating (14) reveals that  $\partial \tilde{B}_N(\bar{\alpha}, \bar{y}) / \partial I(\bar{\alpha}, \bar{y}) < 0$ . When the information gap between current owners and potential buyers widens, prices and, thus, the top owner-determined car quality offered in the market segment must decrease. For a given value of  $\theta$ , this implies that equilibrium prices  $\tilde{P}_N(\bar{\alpha}, \bar{y}) > 0$  inducing  $\tilde{B}_N(\bar{\alpha}, \bar{y}) \geq 0$  may not exist for all market segments  $(\bar{\alpha}, \bar{y})$ . Trade will not occur in segments characterized by a large maximum number of available quality signals  $I(\bar{\alpha}, \bar{y})$ . In such cases, the informational gap between car owners and potential buyers prevents bilateral beneficial trade in such cars.

Next, consider only  $(\bar{\alpha}, \bar{y})$ -segments involving positive trade in the internet. Assume that there exists a parallel market based on the traditional classified ads in print media. Then, if trade occurs in the  $(\bar{\alpha}, \bar{y})$ -segment of this market, the following must clearly hold:

$$\tilde{P}_N(\bar{\alpha}, \bar{y}) - (1 - \gamma_N)c_N J_N^* = \tilde{P}_M(\bar{\alpha}, \bar{y}) - (1 - \gamma_M)c_M J_M^* \quad (16)$$

In order to ensure positive supplies in both markets, the net rewards associated with a sale must be identical. Else, one of the two markets will attract all current owners in this segment. Inserting (16) into (8) then yields  $\tilde{B}_N(\bar{\alpha}, \bar{y}) = \tilde{B}_M(\bar{\alpha}, \bar{y}) \equiv \tilde{B}(\bar{\alpha}, \bar{y})$ . Identical rewards obviously imply an identical top-quality owner-type offering his cars. Utilizing this result and inserting from (8) in (10) thus yields

$$U_M^*(\bar{\alpha}, \bar{y}) = -V_M^*(\bar{\alpha}, \bar{y}) \int_{b_M(\bar{\alpha}, \bar{y})}^{\tilde{B}(\bar{\alpha}, \bar{y})} \exp\{-r\beta\} [\tilde{B}(\bar{\alpha}, \bar{y}) - b_M(\bar{\alpha}, \bar{y})]^{-1} d\beta \quad (17)$$

with

$$V_M^*(\bar{\alpha}, \bar{y}) = \exp\{-r[\theta - \tilde{B}(\bar{\alpha}, \bar{y}) - c_M J_M^* + \frac{1}{2}r \left( \left[ \frac{1}{\sigma^2} + \frac{I(\bar{\alpha}, \bar{y})}{\zeta^2} \right]^{-1} - \left[ \frac{1}{\sigma^2} + \frac{J_M^*}{\zeta^2} \right]^{-1} \right)\} \quad (18)$$

Again, the risk differential in the second line of (18) is associated with the information gap between owners and buyers. It is easily verified that (17) is monotonically

increasing in  $b_M(\bar{\alpha}, \bar{y})$ . While (16) ensures that current owners are indifferent between offering their cars in either market,

$$U_M^*(\bar{\alpha}, \bar{y}) = U_N^*(\bar{\alpha}, \bar{y}) = -1 \quad (19)$$

must then additionally hold for all parallel  $(\bar{\alpha}, \bar{y})$ -segments in order to attract potential buyers. The first equation implies

$$\begin{aligned} & \exp \left\{ -r \left[ c_N J_N^* - c_M J_M^* \right. \right. \\ & \quad \left. \left. + \frac{1}{2} r \left( \left[ \frac{1}{\sigma^2} + \frac{J_N^*}{\zeta^2} \right]^{-1} - \left[ \frac{1}{\sigma^2} + \frac{J_M^*}{\zeta^2} \right]^{-1} \right) \right] \right\} \\ = & \frac{\int_{b_M(\bar{\alpha}, \bar{y})}^{\tilde{B}(\bar{\alpha}, \bar{y})} \exp\{-r\beta\} [\tilde{B}(\bar{\alpha}, \bar{y}) - b_M(\bar{\alpha}, \bar{y})]^{-1} d\beta}{\int_0^{\tilde{B}(\bar{\alpha}, \bar{y})} \exp\{-r\beta\} [\tilde{B}(\bar{\alpha}, \bar{y})]^{-1} d\beta} \end{aligned} \quad (20)$$

Note that the second line on the LHS of (20) now contains the risk differential induced by the information gap between potential buyers in the internet and the traditional classified ad market, respectively. It reflects the different efficient information gathering activity levels in the two markets. The expected utility effect of increasing  $b_M(\bar{\alpha}, \bar{y})$  above zero in the conventional print media market implies that (20) may be satisfied despite this information gap. In this case, let the equilibrium lowest value of owner-determined quality traded in the conventional market be denoted  $\tilde{b}_M(\bar{\alpha}, \bar{y}) > 0$ . Note that the LHS is constant for all such parallel market segments involving positive trade. This follows from the assumption of linear Bayesian learning combined with constant unit information costs. Then, differentiating the RHS reveals  $d\tilde{B}(\bar{\alpha}, \bar{y})/db_M(\bar{\alpha}, \bar{y}) < 0$ .

Thus, as the information gap between potential buyers and current owners widens - yielding  $\partial \tilde{B}(\bar{\alpha}, \bar{y})/I(\bar{\alpha}, \bar{y}) < 0$  as discussed above - the difference  $[\tilde{B}(\bar{\alpha}, \bar{y}) - \tilde{b}_M(\bar{\alpha}, \bar{y})]$  decreases. Yet, it cannot decrease beyond  $[\tilde{B}(\bar{\alpha}, \bar{y}) - \tilde{b}_M(\bar{\alpha}, \bar{y})] = 0$  - due to (8) and (16). Otherwise all current owners would again move to only one market. Hence, parallel market segments with positive trade must be associated with relatively low values of  $I(\bar{\alpha}, \bar{y})$ . The cars offered in such segments are characterized by relatively few additional signals concerning their technical reliability in the owner's information set. As more of such signals become available and known to current owners - thus, the increase in  $I(\bar{\alpha}, \bar{y})$  widens the information gap between current owners and potential buyers -  $[\tilde{B}(\bar{\alpha}, \bar{y}) - \tilde{b}_M(\bar{\alpha}, \bar{y})]$  converges to zero. Even higher values of  $I(\bar{\alpha}, \bar{y})$  then imply that such  $(\bar{\alpha}, \bar{y})$ -segments only exist in the internet market. Due to lower information



costs, this market generally entails a smaller information gap between current owners and potential buyers. As shown above, for extreme  $I(\bar{\alpha}, \bar{\gamma})$ -values positive trade may not be possible at all. Such segments cannot be found in either the internet or the conventional print media market.

## 2.4 Parallel markets and testable hypotheses

Adverse selection equilibria exist due to the inobservability of quality signals in markets for experience goods. This obviously implies that - despite the "lemon" model's frequent inclusion in modern micro-economic textbooks - direct empirical tests can hardly be conducted. It appears implausible that the researcher is able to collect quality signals which cannot be observed by the agents in the market. However, Bond (1982) uses a maintenance cost survey on pick-up trucks which provides *ex-post* quality information. The study compares trucks bought new, respectively used by their current owners and controls for quality differences which can be communicated between buyers and sellers - hence, the  $\bar{\gamma}$ -characteristics introduced above. Significant differences in the probabilities to require major repairs cannot be confirmed. Thus, trade in used pick-up trucks does not appear to induce an unobservable seller-contingent quality risk. The fact that used pick-ups did require more maintenance on average is attributed to "*significantly higher lifetime mileage in each model year*". This is seen to constitute "bad news" for potential buyers. As noted above, mileage per year may not constitute a monotone quality signal, however. Further, the analysis does not control for differences in brands and accessory units. Also, the pick-ups in the sample are predominantly used commercially. The respective buyers and sellers are therefore not typical agents in a "lemon" market.

In contrast, the current empirical investigation uses data on traditional and e-commerce trade in used "family" cars. The occurrence of equilibrium parallel trade - in cars characterized by identical manufacturer specifications and advertised quality information - then provides an indirect test of the adverse selection hypothesis. Note that the equilibrium is in fact incentive-compatible. Obviously, the net rewards for current owners are equalized. Thus, potential sellers of  $(\bar{\alpha}, \bar{\gamma})$ -cars randomly select over the internet and the traditional classified ad market, if this segment exists in both markets. At the same time, potential buyers are indifferent with regard to shopping over all segments and markets. Hence, they also randomly select. Now, suppose that the buyers' information gathering activities are less costly in the internet. Then, only

if additional seller-contingent quality risk induces adverse selection, the internet will not strictly dominate the traditional market-place. In equilibrium the relatively higher costs of information gathering must be exactly compensated by less adverse selection in the traditional market. Two different sources of inefficiency are thus necessary in order to generate positive equilibrium trade in parallel market segments. In fact, this constitutes a standard result of second-best analysis [Lipsey and Lancaster (1956)]. Although the test of existing adverse selection is indirect, it necessarily only requires the assumption of competitive equilibrium between e-commerce and traditional trade. Disputing this assumption implies that parallel market places must reflect agent heterogeneity with respect to individual market-entrance costs. Consequently, market-entry decisions would be agent-specific and not contingent on advertised information signalling the attrition-state of the car.

However, some segments may only exist in the market environment exhibiting lower information gathering costs for potential buyers. Very old cars may be considered as "oldtimers". These are highly valued by their individual owners and - if at all - sold in specific markets. More commonly, however, cars age until they possess only "scrap value". In part, the existence of scrap yards certainly reflects the elimination of trade due to the information gap between potential buyers and sellers with respect to technical quality risk. Consequently, all used cars which are not scrapped must be traded in the market environment involving lower information gathering costs. The high information cost market will then lack some segments which - according the observable signals  $(\bar{\alpha}, \bar{y})$  - imply low quality. Thus, the assertion that internet trade is associated with lower information costs implies that such  $(\bar{\alpha}, \bar{y})$ -cars can be found in the internet, while not being advertised in the conventional classified ad market. Comparing the offer distributions between the two markets therefore allows to test the information dominance assertion for electronic trade without referring to prices.

Despite lower search costs, additional mark-ups and risks associated with new intermediation services [OECD (1997, p. 9-10), Zerdick, et al. (1999, p. 149-154)] and the necessity of intensified advertising in the pursuit of differentiation strategies [OECD (1999, p. 73-75), Choi, et al. (1997, p. 346-367)] render price effects generally ambiguous. Applied to the classified ad market for used cars these arguments appear inadequate, however. First, the informational asymmetry is ultimately associated with an individual seller's personal characteristics in a spot-market environment. Second, the information communicated via classified ads is in fact identical in the two market-

places. Hence, lower information costs must be associated with lower prices unless consumers in the internet incur a smaller share of these costs. Yet, the e-commerce literature unanimously agrees that consumer search activities in the internet are more intense. Obviously, the development of "shopbots" and "chat groups" in which individual experiences can be communicated reflect an increased search activity level of the internet consumer. Empirically, Hoque and Lohse (1999) further support this assertion.

Obviously, the characteristic model properties mirror the arguments found in the current descriptive literature on e-commerce. In particular, the model captures both the effects of search cost dominance and of a more severe selection problem associated with the new electronic marketplace. However, these views are not opposing - yielding either optimistic or rather sceptical scenarios concerning the future development of internet trade. In equilibrium both markets may co-exist. Such parallel markets can even continue to exist, if potential buyers augment the signal vector communicated in the traditional classified ad market by engaging in additional internet research. Only if the distribution of information costs, the information gathering means, and the respective activity levels are identically equal in both markets, trade must occur in a single market - characterized by identical prices for all  $(\bar{\alpha}, \bar{y})$ -cars advertised.

### 3 An explorative empirical investigation

#### 3.1 Data description

The empirical analysis focuses on internet and print media classified ads for Volkswagen Rabbits series III and BMW series 3xx. This brand and type combination has been chosen for two reasons. First, both vehicle types are very widespread in Germany. This allows to collect a sufficient number of observations in both marketplaces. More importantly, however, the general public as well as the automobile experts' interest in both brand models can be assumed to be equally developed in Germany. Also, VW and BMW representatives exist in virtually every German city. Second, the two model types do not constitute close substitutes from the point of view of potential buyers. BMWs rather appeal to auto-sports interested customers, while the VW Rabbit is a typical economy car. Including the brand variable in the analysis can therefore provide additional evidence with respect to model specification. The respective vectors of

additional signals - denoted  $\bar{z}^o$  in the analysis above - should possess approximately equal cardinality. Hence, if agents randomly select over marketplaces, this variable should not attain statistical significance in explaining differences in the probabilities of being offered in a particular market. At the same time, brand name should prove to exhibit price effects reflecting consumer preferences.

Obviously, observations cannot be drawn from a complete, or representative sample of used cars traded in Germany. Moreover, the theoretical model rests on the existence of unobserved selection in the offer distributions. Thus, standard techniques to control for sample selection bias cannot be employed sensefully. It was therefore decided to minimize this problem by clustering the data sources. The researchers defined a number of regional and national newspapers to be searched for the respective classified ads. Internet pages were selected randomly. Further, the data collection was confined to classified ads which appeared in the second week of June 1999. Undergraduate students then compiled complete sets of offer observations within these pre-specified media and time-span. Nevertheless, the results reported in the following may still be subject to sample selection bias. Rather than providing a representative analysis, the current empirical investigation is therefore explorative in nature.

Following this procedure, the empirical investigation can be based on a total of 449 observations for the two brands. Among these, 271 have been derived from newspapers. The data set contains 45 % BMW-observations in the traditional print media and 25 % respective internet entries. In the following regression analyses VW (BMW) then attains the value 1 (0). In addition to brand name, model type, and the price requested, the ads contain information concerning air conditioning, sun-roofs, air-bags for both front passengers, size of engine in 1000 cm<sup>3</sup>, and horse power. Also, all ads named the date of first licensing and provided information concerning the total kilometers (km) driven. Kilometers per month have then been calculated as an additional quality signal - completing the vector  $\bar{y}$  as discussed above. Table 1 provides the list of available advertisement information and the respective abbreviations used in the following. It also reports some descriptive statistics.

### **3.2 Differences in the quality distributions**

The assumption of internet dominance with respect to the information gathering costs implies that - in competitive equilibrium - e-commerce trade occurs over all segments, while the traditional market will lack some low-quality segments. The descriptive

statistics of *Table 1* already provide some first hints concerning such signal-contingent market-entry decisions of sellers. Comparing internet and traditional marketplaces, minimum ages and total kilometers appear to exhibit less pronounced differences than the respective maximum values. Moreover, it is possible to perform t-tests on statistically significant differences in means, Wilcoxon-Mann-Whitney tests on median values, and 3F-tests on variances for these three variables. For both total kilometers driven and age all tests strongly confirm the existence of significant differences. The rejection probabilities are below 0.001. Kernel-density estimations can further be employed in order to visualize the differences in the distributions. Following Pagan and Ullah (1999, p. 23-29), Epanechnikov kernels have been calculated for total kilometers and ages. The graphical representations of the respective density estimations in *Exhibits 1* and *2* then illustrate the differences in the distributions of these two observable quality signals in the two marketplaces.

Thus, on average e-commerce entails trade in older cars with more total kilometers driven. Both clearly represent "bad news" for potential buyers. For obvious reasons, such significant differences cannot be confirmed for kilometers per month. In fact, only the respective 3F-test on variances attains a significance level of 0.06. However, as discussed above, kilometers per month - utilized by Bond (1982) - may not constitute a monotone quality signal. Also, the first five variables listed in *Table 1* - associated with the  $\bar{\alpha}$ -vector entries - can only imperfectly be associated with the occurrence of signal-contingent market-entry decisions. The manufacturer's model specifications usually preclude choices with respect to combinations of horse-power and engine size. Similarly, in older models air-conditioning usually requires a larger engine size. Moreover, airbags for both front passengers and air-conditioning have become standard in more recent model years. Respective quality signals - applicable to older models only then - have therefore be constructed by weighting these binary variables by age. These variables enter the regression analysis carried out below. However, mean, median, and variance tests on such combined variables cannot be interpreted sensefully and have therefore been omitted.

*Table 2a* then reports the results of a Probit analysis estimating the probability of observing an internet, rather than a newspaper classified ad for a particular used car. The advertisement costs increase with more detailed descriptions of the car. Hence, economic rationality implies that all information communicated via such ads should enter the potential buyers' calculations of informative signals. Yet, from a research

point of view it cannot be assessed how such calculations are carried out by the agents. Hence, the regression equation reflects a naive approach. Basically, it includes all information found in the classified ads. In only one respect, 'goodness-of-fit' considerations have entered the model specification. As expected, age and total kilometers appear to be very significant in determining the selection effect. Both variables are also highly correlated. They enter the regression equation a second time when calculating kilometers per month. Alternative regressions have then shown that additionally including the square of age significantly improves the model's explanatory power.

Obviously, the Probit-analysis generally confirms the effects of quality-contingent market-entry decisions as well. The advertisement information exhibits a statistically significant impact on the probability of observing an internet offer. Also, all signs agree with the expectations derived from the theoretical model and the respective stylized fact discussion concerning e-commerce. Internet trade selects lower quality segments defined by the advertised signals. Among these signals, only the air-condition variable remains insignificant. As discussed above, the respective signalling effect may to a considerable extent already be captured by engine size and brand name. Interestingly, brand name itself - captured by the variable VW - does not exhibit a significant impact. As discussed above, this should be expected, if the sets of available additional information for the two car types possess identical cardinality. From a more practical point of view, it indicates that the problem of selection bias introduced by the data collection process is not severe. *Table 2b* also reports the results of two standard model specification tests for Probit-regression analysis [Andrews (1988), Hosmer and Lemeshow (1989)]. Both strongly confirm that the null-hypothesis of misspecification must be rejected.

### **3.3 E-commerce price effects**

Obviously, the current investigation cannot be based on actual exchange prices. Thus, the following discussion requires the assumption of sufficient correlation between advertised and realized prices. *Exhibit 3* then depicts Epanechnikov kernel estimates for the price densities. They visualize the existing differences in price dispersion between the two marketplaces. Again, the mean value t-test, the Wilcoxon-Mann-Whitney test on median values, and the 3F-test on variances confirm statistically significant differences. The respective rejection probabilities are all below 0.001. Further, the price variable has been inserted into Probit estimates concerning the probability to find a particular offer in the internet. In this case, it attains very dominant explanatory power - to

a considerable extent crowding out the effects associated with the quality variables contained in *Table 2a*. However, given the results above, such estimates do not allow plausible interpretations. The respective regressions have therefore been omitted.

From the point of view of empirical analysis, this poses a rather severe model specification problem, however. Obviously, the agents utilize the same information in forming their price expectations which also governs the selection of sellers over marketplaces. Again, the particular functional form relating advertized signals to expectations is unknown. Thus, *Tables 3a-b* and *Table 4* report price estimates for the two marketplaces separately and for the pooled data set. In each case the list of explanatory variables has been chosen identically equal to the one included in the Probit analysis above. Clearly, the OLS-regressions perform rather well. This highlights the problem of specifying a price equation which at the same time appropriately adjusts for the seller selection effect over the two marketplaces.

The model specification tests concerning the Probit-analysis have already confirmed signal-contingent market-entry decisions. Thus, the OLS-approach of *Table 4* must obviously be rejected. Instead, the marketplace variable "internet" has to be instrumentalized. The respective two-stage least squares approach (Two-SLS) can then be found in *Table 5*. The particular structural form of the equations has been chosen rather naively again. Note that the "internet" variable is instrumentalized using the functional relationship which has performed best in the Probit-analysis. The price equation then includes the observable quality variables which possess statistical significance in the respective OLS-regressions. The only exception is  $[\text{Age}]^2$ . Omitting this variable from the price equation raises the model's explanatory power and, at the same time, the significance level of the marketplace variable. Including the brand type and the air-conditioning variables - which are insignificant in the Probit-analysis - in the price equation has contributed to model performance in the same way.

It should be noted that comparisons of the  $R^2$  between OLS and Two-SLS regressions are generally invalid. Yet, the results reported in *Table 5* reveal that the Two-SLS approach performs rather well. In particular, the significance level of the internet variable increases. Given the distributional and Probit model specification tests above, this constitutes a desirable feature of an appropriate price estimate. Moreover, instrumentalizing "internet" increases the respective coefficient rather drastically compared to the OLS-approach of *Table 4*. Again, this is consistent with the expectation derived from the theoretical model. It predicts that e-commerce should be associated with two

separate price effects. First, the used cars traded electronically exhibit lower average quality according to the advertised signals. Simultaneously, lower information costs and a higher level of information gathering activities by potential buyers contributes a second tendency towards lower prices. Thus, appropriately accounting for signal-contingent market-entry decisions introduces a negative price-effect which adds to the simultaneously existing quality-price relationship in the specific market.

However, the estimated DM 4,312 price differential between electronic and traditional classified ads then also reflects a combined price effect. Given the lack of knowledge concerning the conditional quality expectations formed by potential buyers, the two effects cannot be separated empirically. Nevertheless, comparing the coefficients of the internet variable in the OLS and the Two-SLS regressions suggests that price differentials between the electronic and the traditional marketplace to a considerable extent reflect quality selection. This contrasts with previous studies emphasizing the efficiency-enhancing virtue of internet trade.

### **3.4 Summary and conclusions**

The emergence and rapid growth of e-commerce is commonly seen to reflect the advantages of a dominant information environment. Hence, it potentially implies a substitution of traditional marketplaces. Yet, if seller-contingent quality uncertainty induces persistent adverse selection, a competitive equilibrium with parallel segments in both electronic and traditional marketplaces arises. The migration of sellers induces more adverse selection in internet trade which exactly compensates the lower costs of information gathering activities. Market segments are defined conditional on quality signals communicated via the product advertisements. Given that the information costs generally preclude positive trade in very low quality segments, some adjacent segments further exist only in the internet. Accepting the assumption of competitive equilibrium, this allows for an indirect test of the adverse selection hypothesis. Market-entry decisions must be contingent on the advertised information rather than reflecting a distribution of agent-specific transaction costs.

The empirical investigation of German classified ads for used cars generally supports this assertion. Although it lacks a representative data source, the analysis strongly confirms the effects of signal-contingent market-entry decisions by potential sellers. This follows from statistical tests on differences in the offer distributions as well as from the subsequent Probit-regression and the respective model specification tests. These results



then imply a Two-SLS approach for estimating the price equation. The estimated price reduction of DM 4312 associated with internet trade warrants some qualifications, however. First, it applies to the current sample only. Second, the estimate is based on rather naive "goodness-of-fit" consideration, since the true functional form relating advertised quality signals to expectations remains unknown. Finally, two price effects which in theory contribute simultaneously cannot be separated empirically. On the one hand, lower e-commerce prices signal a higher propensity of observing low-quality offers conditional on the advertised information. At the same time, prices will additionally be decreased due to lower search costs and a higher information gathering activity level of internet buyers.

Hence, the study contributes to the debate on the future of e-commerce and traditional marketplaces. In particular, it has been shown that the two markets may co-exist despite dominant search means available in the internet. Moreover, this conclusion does not require that e-commerce generally induces a more severe information problem with respect to seller characteristics. A compensatory migration of potential sellers between marketplaces suffices to support the equilibrium. The implications for strategic market-entry decisions of firms are straightforward then. They particularly apply to retailers, intermediaries, and providers of other services in which the buyers' quality expectations appreciate product as well as seller characteristics. Entering the electronic marketplace, firms face a more severe adverse selection problem. This particularly affects high-quality producers. These must therefore increase their efforts in communicating reliable seller information to potential customers. Hence, a high-quality producer-seller must balance the possible benefits of market extension by participating in e-commerce against the higher communication costs necessitated by this marketplace. Some authors have suggested that network economies actually reinforce the adverse selection problem. In this case, the current results clearly obtain even greater significance.

## 4 References

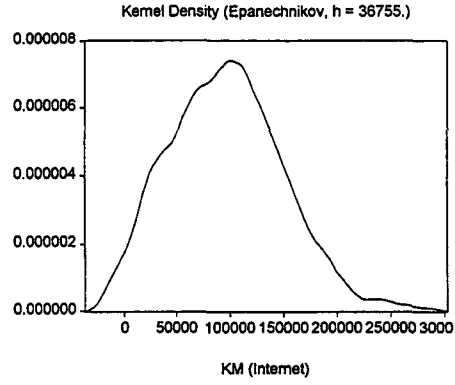
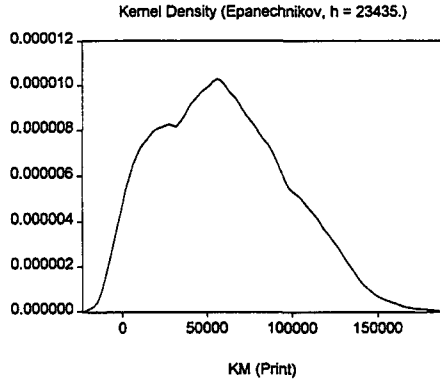
- Akerlof, G. A. (1970) The Market for 'Lemons': Quality Uncertainty and the Market Mechanism. *Quarterly Journal of Economics*, vol. 89, p. 345-364.
- Alba, J., J. Lynch, B. Weitz, C. Janiszewski, R. Lutz, A. Sawyer and S. Wood (1997) Interactive Home Shopping: Consumer, Retailer, and Manufacturer Incentives to

- Participate in Electronic Marketplaces. *Journal of Marketing*, vol. 61, p. 39-52.
- Andrews, D. W. K. (1988) Chi-Square-Diagnostic Tests for Econometric Models: Introduction and Application, *Journal of Econometrics*, vol. 37, p. 135-156.
- Bakos, J. Y. (1997) Reducing Buyer Search Costs: Implications for Electronic Marketplaces. *Management Science*, vol. 43, p. 1676-1692.
- Barua, A., J. Pinnell, J. Shutter, and A.B. Whinston (1999) *Measuring the Internet Economy: An Exploratory Study*. Austin, TX: Center for Research in Electronic Commerce, University of Texas, download from: <http://www.crec.bus.utexas.edu>.
- Bond, E. W. (1982) A Direct Test of the "Lemons" Model: The Market for Used Pickup Trucks. *American Economic Review*, vol. 72, p. 836-840.
- Campbell, D.E. (1995) *Incentives: Motivation and the Economics of Information*. Cambridge, England, etc: Cambridge University Press.
- Choi, S.-Y., D.O. Stahl and A.B. Whinston (1997) *The Economics of Electronic Commerce*. Indianapolis, In: McMillan Technical Publishing.
- De Groot, H.M. (1970) *Optimal Statistical Decisions*. New York, etc.: McGraw-Hill.
- Gemini (1999) *Content and Commerce Driven Strategies in Global Networks - CONDINET*. Issy-Les-Moulineaux, France: Gemini Consulting, and Luxembourg: European Commission, Directorate General XIII/E, executive summary, download from: <http://www2.echo.lu/condrinet>.
- Greenwald, A.R. and J.O. Kephart (1999) Shopbots and Pricebots. *Proceedings of Sixteenth International Joint Conference on Artificial Intelligence*, forthcoming, download from: <http://www.research.ibm.com/infoecon/researchpapers.html>.
- Greenwald, A.R., J.O. Kephart and G.J. Geraldo (1999) Strategic Pricebot Dynamics. *Proceedings of ACM Conference on Electronic Commerce*, forthcoming, download from: <http://www.research.ibm.com/infoecon/researchpapers.html>.
- Hoffman, D.L. and T.P. Novak (1996) Marketing in Hypermedia Computer Mediated Environments. *Journal of Marketing*, vol. 39, p. 50-68.

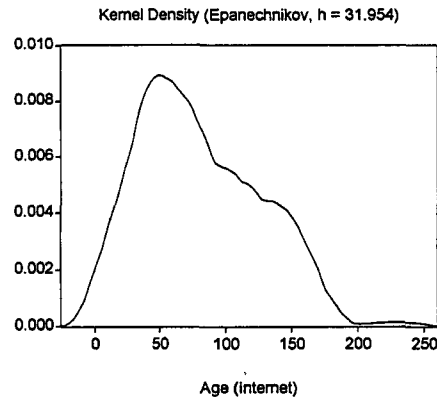
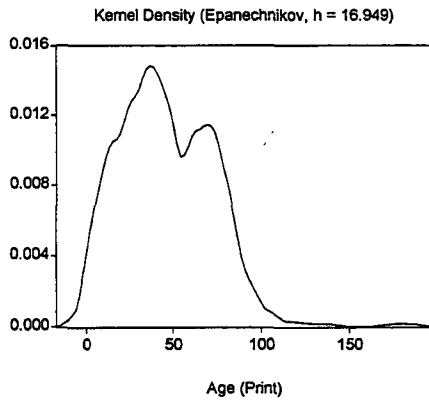
- Hoque, A.Y. and G.L. Lohse (1999) An Information Search Cost Perspective for Designing Interfaces for Electronic Commerce. *Journal of Marketing Research*, vol. 26, p. 387-394.
- Hosmer, D. W. J. and S. Lemeshow (1989) *Applied Logistic Regression*, Chichester, England, and New York: John Wiley & Sons.
- Internet Economy Indicators (1999) *October 1999 Internet Economy Indicators*. Austin, TX: Center for Research on Electronic Commerce, University of Texas, download from: <http://www.internetindicators.com>.
- Kahin, B. and J. Keller (1995) *Public Access to the Internet*. Cambridge, MA, and London, England: MIT Press.
- Kalakota, R. and A.B. Whinston (1999) *Do or Die: Market Segmentation and Product Positioning on the Internet*. Austin, TX: Center for Economic Research in Electronic Commerce, University of Texas, available as: <http://www.cism.bus.utexas.edu/res/articles/segmentation.html>.
- Keeney, R.L. (1999) The Value of Internet Commerce to the Customer. *Management Science*, vol. 45, p. 533-542.
- Kephart, J.O. and A.R. Greenwald (1999) Shopbot Economics. *Proceedings of the Fifth European Conference on Symbolic and Quantitative Approaches to Reasoning under Uncertainty*, forthcoming, download from: <http://www.research.ibm.com/infoecon/researchpapers.html>.
- Kirby, A.J. (1993) Optimal Information Exchange. *Information Economics*, vol. 5, p. 5-29.
- Lipsey, R.G. and K.J. Lancaster (1956) The General Theory of the Second-Best. *Review of Economic Studies*, vol. 24, p. 11-32.
- McKnight, L.W. and J.P. Baily (1997) *Internet Economics*. Cambridge, MA, and London, England: MIT Press.
- Mesenbourg, T.L. (1999) *Measuring Electronic Business: Definitions, Underlying Concepts, and Measurement Plans*. Washington, D.C.: US Bureau of Census, available as: <http://www.ecommerce.gov/ecomnews/e-def.html>.

- Milgrom, P.R. and J. Roberts (1992) *Economics, Organization and Management*. Englewood Cliffs, N.J.: Prentice Hall.
- OECD (1997) *Business-to-Consumer Electronic Commerce: Survey of Status and Issues*. Paris, France: Head of Publication Services, OECD, document no. OCDE/GD(97) 219, download from: <http://www.oecd.org/dsti/sti/it/ec>.
- OECD (1999a) *Defining and Measuring E-Commerce: A Status Report*. Paris, France: Head of Publication Services, OECD, document no. DSTI/ICCP/IIS(99)4/FINAL, download from: <http://www.oecd.org/dsti/sti/it/ec>.
- OECD (1999b) *The Economic and Social Impact of Electronic Commerce: Preliminary Findings and Research Agenda*. Paris, France: Head of Publication Services, OECD, OECD code 9319990011P1, download from: <http://www.oecd.org/dsti/sti/it/ec>.
- Pagan, A. and A. Ullah (1999) *Non-Parametric Econometrics*, Cambridge, England: Cambridge University Press.
- Peterson, R.A., S. Balasubramanian and B.J. Bronneberg (1997) Exploring the Implications of the Internet on Consumer Marketing. *Journal of the Academy of Marketing Science*, vol. 25, p. 329-346.
- US Department of Commerce (1999) *The Emerging Internet Economy II*. Washington, D.C.: Secretariat on Electronic Commerce, US Department of Commerce, download from: <http://www.ecommerce.gov>.
- Zerdick, A., A. Picot, K. Schrape, A. Artropé, K. Goldhammer, U.T. Lange, E. Vierkant, E. López-Escobar and R. Silverstone (1999): *Die Internet-Ökonomie: Strategien für die digitale Wirtschaft*. Berlin, etc.: European Communication Council Report, Springer (English version entitled "E-conomics: The Economy of the E-Commerce and the Internet" forthcoming).

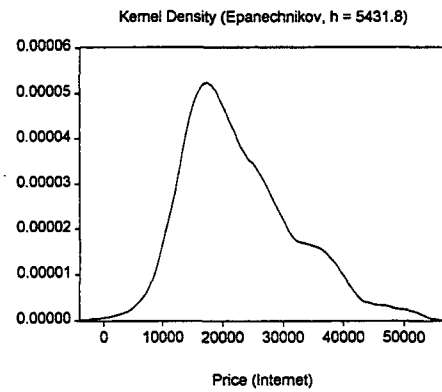
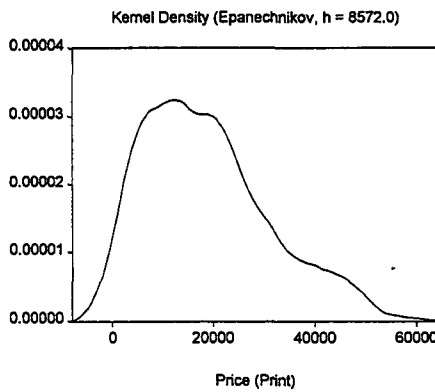
### Exhibit 1: Kernel-Density estimation (Epanechnikov): Kilometers



### Exhibit 2: Kernel-Density estimation (Epanechnikov): Age



### Exhibit 3: Kernel-Density estimation (Epanechnikov): Prices



**Table 1: Descriptive sample statistics**

Market-Place	Print Media		Internet	
Variable	Mean (Median)	Maximum (Minimum)	Mean (Median)	Maximum (Minimum)
Sun-roof [SR]	.4269 (n.a.)	1 (0)	.3321 (n.a.)	1 (0)
Airbags [AB]	.32 (n.a.)	1 (0)	.39 (n.a.)	1 (0)
Air-Conditioning [AC]	.32 (n.a.)	1 (0)	.24 (n.a.)	1 (0)
Engine Size [ENG]	1.7 (1.6)	2.8 (1.4)	1.8 (1.8)	3.5 (1.4)
Horse Power [HP]	96 (101)	402 (60)	116 (112)	250 (60)
Months of age [AGE]	46.2 (42.0)	180.0 (1.0)	71.9 (61.5)	228.0 (6.0)
Total Km [KM]	59,212 (56,000)	165,000 (100)	92,585 (85,500)	265,000 (200)
Km per month [MKM= KM/AGE]	1,339 (1,225)	8,929 (100)	1,408 (1,290)	5,294 (25)
Price [P]	22,498 (20,500)	51,450 (1,500)	20,168 (18,500)	55,500 (700)

**Table 2a: Estimate of the probability to observe an internet offer**

ML-Probit* - Dependent Variable:	Internet Offer =1	Included: 439 observations	Internet: 168 observations Print Media: 271 observations	
Variable	Coefficient	Std. Error	z-Statistic	Prob.
AB*AGE	- 0.018585	0.003939	-4.717660	0.0000
AC*AGE	-0.002019	0.003868	-0.521834	0.6018
SR	-0.411067	0.143809	-2.858422	0.0043
AGE	0.032904	0.012077	2.724577	0.0064
[AGE] <sup>2</sup>	-0.000235	8.35exp{-5}	-2.813589	0.0049
ENG	1.348965	0.421945	3.197017	0.0014
VW	0.085048	0.205292	0.414278	0.6787
HP	-0.018009	0.005450	-3.304663	0.0010
KM	-1.76exp{-5}	4.64exp{-6}	-3.785512	0.0002
MKM	0.000480	0.000192	2.506597	0.0122
Mean dependent var.:	0.615034	S.D. dependent var.:	0.487142	
S.E. of regression:	0.401390	Akaike info criterion:	1.009389	
Sum squared errors:	69.11794	Schwarz criterion:	1.102430	
Log likelihood:	- 211.5610	Hannan-Quinn crit.:	1.046098	
Avg. log likelihood:	- 0.481916			

Notes: \*) Convergence after 11 iterations; 9 observations excluded after adjusting end-points; QML (Huber/White) standard errors and covariance.

**Table 2b: Specification tests for the Probit-model**

Andrews and Hosmer-Lemshow Goodness-of-Fit Tests <sup>+</sup> )								
	Quantile	of Risk	Dep. =0	Dep. =1	Total	H-L		
	Low	High	Acual	Expect.	Actual	Expect.	Obs.	Value
1	2exp{-11}	0.1439	41	40.7100	2	2.29002	43	0.03880
2	0.1441	0.3585	30	33.0669	14	10.9331	44	1.14475
3	0.3586	0.4949	31	25.0251	13	18.9749	44	3.30791
4	0.4981	0.6137	20	19.2257	24	24.7743	44	0.05538
5	0.6149	0.6870	14	15.4861	30	28.5139	44	0.22006
6	0.6878	0.7566	11	12.3953	33	31.6047	44	0.21866
7	0.7566	0.8064	13	9.54409	31	34.4559	44	1.59801
8	0.8064	0.8541	5	7.59613	39	36.4039	44	1.07242
9	0.8552	0.9291	3	4.81358	41	39.1864	44	0.76723
10	0.9306	1.000	1	1.98693	43	42.0131	44	0.51340
Totals:			169	169.850	270	269.150	439	8.93663
H-L Statistic: 8.9366    Andrews Statistic: 9.9223								
Prob. $\chi^2(8)$ : 0.3477    Prob. $\chi^2(10)$ : 0.4437								

Note: <sup>+</sup>) Groupings based on predicted risk.



Table 3a: OLS-estimates of price function for the internet market

Dependent variable $P^*$ )				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
Constant	23814.42	2057.942	11.57196	0.0000
AB*AGE	-3.526029	14.15479	-0.249105	0.8035
AC*AGE	21.87804	12.55736	1.742248	0.0827
SR	753.0791	713.0661	1.056114	0.2919
AGE	-268.8464	40.08796	-6.706414	0.0000
[AGE] <sup>2</sup>	0.606907	0.158323	3.833341	0.0002
ENG	-71.15683	208.8139	-0.340767	0.7336
VW	-2601.948	640.0716	-4.065090	0.0001
HP	145.1551	16.20514	8.957347	0.0000
KM	-0.050649	0.019263	-2.629342	0.0091
MKM	-0.269319	0.783644	-0.343675	0.7314
$R^2 = 0.794067$ Adj. $R^2 = 0.786116$ Prob(F-Stat.) = 0.000000 S.E. of regr.: 4099.308    Sum squ. res.: $4.35 \exp\{9\}$ F-stat.: 99.86886 Lg. likelih.: -2623.513    Durb.-Wats. stat.: 1.876771    Mean $P$ : 22534.83 Akaike crit.: 19.51491    Schwarz crit.: 19.66151    S.D. $P$ : 8863.821				

Notes: \*) Included observations 270 after adjusting for endpoints; White heteroskedasticity-consistent standard errors and covariances.

Table 3b: OLS-estimates of price function for the traditional market

Dependent variable $P^+$ )				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
Constant	35412.04	2382.879	14.86103	0.0000
AB*AGE	-20.12771	16.28900	-1.235663	0.2184
AC*AGE	-1.339782	13.35693	-0.100306	0.9202
SR	686.0527	642.4875	1.067807	0.2872
AGE	-377.6153	36.70719	-10.28723	0.0000
[AGE] <sup>2</sup>	0.930644	0.142203	6.544469	0.0000
ENG	1618.761	2319.928	0.697764	0.4864
VW	-5926.910	1182.226	-5.013347	0.0000
HP	67.82212	27.85429	2.434889	0.0160
KM	-0.044161	0.013095	-3.372344	0.0009
MKM	0.100731	0.968089	0.104052	0.9173
R <sup>2</sup> =0.918521    Adj. R <sup>2</sup> = 0.913364    Prob(F-Stat.)=0.000000				
S.E. of regr.: 3548.781    Sum squ. res.: 1.99exp{9}    F-stat.: 178.1141				
Lg. likelih.: -1615.580    Durb.-Wats. stat.: 1.595161    Mean $P$ : 18678.63				
Akaike crit.: 19.24947    Schwarz crit.: 19.45319    S.D. $P$ : 12056.72				

Notes: +) Included observations 169 after adjusting for endpoints; White heteroskedasticity-consistent standard errors and covariances.

Table 4: OLS-estimate of price function over both marketplaces

Dependent variable $P^*$ )				
Variable	Coefficient	Std. Error	t-Statistic	Prob
Constant	29189.98	1591.706	18.33881	0.0000
Internet	-974.8213	505.8215	-1.927204	0.0546
AB*AGE	-4.181966	10.60351	-0.394394	0.6935
AC*AGE	4.443650	8.845336	0.502370	0.6157
SR	833.4337	499.2871	1.669247	0.0958
AGE	-284.5998	24.59286	-11.572246	0.0000
[AGE] <sup>2</sup>	0.563638	0.090492	6.228581	0.0000
ENG	-192.9948	381.6412	-0.5005697	0.6133
VW	-3525.224	539.8254	-6.530305	0.0000
HP	113.9753	10.51747	10.836676	0.0000
KM	-0.051004	0.009799	-5.205274	0.0000
MKM	-0.106827	0.529914	-0.201592	0.8403

$R^2 = 0.854069$     Adj.  $R^2 = 0.850310$     Prob(F-Stat.) = .000000  
 S.E. of regr.: 4012.144    Sum squ. res.:  $6.87 \exp\{9\}$     F-stat.: 227.1862  
 Lg. likelih.: -4259.249    Durb.-Wats. stat.: 1.704487    Mean  $P$ : 21050.32  
 Akaike crit.: 19.45899    Schwarz crit.: 19.57064    S.D.  $P$ : 10370.03

\*) Notes: Included observations 439; 9 observations excluded after adjusting endpoints; White heteroskedasticity-consistent standard errors and covariances.

**Table 5: Two-stage least squares estimate of price function over both marketplaces**

Dependent variable $P$ - Pooled data*)				
Internet instrumentalized; instrument list identical to list of dependent variables in Probit-regression of Table 2.				
Variable	Coefficient	Std. Error	t-Statistic	Prob
Constant	31304.71	2479.264	12.62661	0.0000
Internet	-4311.890	1714.403	-2.476101	0.0137
HP	103.7359	11.38773	9.109446	0.0000
AGE	-197.3381	10.22197	-19.30529	0.0000
KM	-0.069708	0.007153	-9.745772	0.0000
VW	-3886.391	552.7171	-6.995244	0.0000
AC*AGE	-6.702464	11.47054	-0.584320	0.5593
$R^2 = 0.827716$ $Adj. R^2 = 0.825323$ $Prob(F-Stat.) = .000000$ S.E. of regr.: 4334.090    Sum squ. res.: $8.11 \exp\{9\}$ F-stat.: 352.2315 Durb.-Wats. stat.: 1.622172    S.D. $P$ : 10370.03    Mean $P$ : 21050.32				

\*) Notes: Included observations 439; 9 observations excluded after adjusting endpoints.