Englmaier, Florian; Schmöller, Arno; Stowasser, Till

Conference Paper

Price Discontinuities in an online used Car Market

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2013: Wettbewerbspolitik und Regulierung in einer globalen Wirtschaftsordnung - Session: Market Theory and Applications, No. E14-V1

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Price Discontinuities in an online used Car Market*

Florian Englmaier† Arno Schmöller‡ Till Stowasser§

March 1, 2013

Abstract

We examine empirically whether individuals evaluating used cars efficiently aggregate all relevant information on its constituent characteristics. Based on detailed field data on more than 80,000 used car offers in a large online marketplace, we provide evidence for biased information processing. While the precise date of first registration, i.e., its "age", is publicly and prominently stated for each car, we identify an amplified value adjustment for otherwise identical cars at year-count changes. These discontinuities indicate that individuals over-react to the figure displayed in the latter, while underrating the finer information on a car's age as conveyed through the month of first registration. Moreover, we are able to replicate the findings from Lacetera *et al.* (2012) and find discontinuous drops in prices at 10,000km odometer thresholds. While the latter finding, as suggested by Lacetera *et al.* (2012), is consistent with a left-digit bias in the processing of numerical information, the first finding cannot be explained by this. Our findings underline that information-processing heuristics matter also in markets with large stakes and easily observed information.

Keywords: Complex Goods; Price Discontinuities; Information Neglect; Heuristics; Field Study

JEL classification: D12, D83, L 62

---

*We are especially grateful to Anton Vasilev for his great support, interest and valuable hints. We thank Matthias Dischinger and Klaus Schmidt for their helpful comments and suggestions. Ines Helm provided excellent research assistance. This research was partially funded through DFG grant *SFB/TR-15*. The paper was previously circulated as: Florian Englmaier and Arno Schmöller (2009) "The Evaluation of Complex Goods: Evidence from Online Car Sales".

†University of Würzburg, florian.englmaier@uni-wuerzburg.de

‡TWS Partners, info@schmoeller.org

§University of Würzburg, till.stowasser@uni-wuerzburg.de

1 Introduction

Modern consumer goods are often complex in nature. They consist of a multitude of characteristics that collectively account for their market value. People are challenged to evaluate goods like cars, mobile phones, or personal computers on the basis of their various constituting attributes. While economic theory suggests that a rational agent should incorporate all relevant pieces of information into his considerations and exclude any that are non-informative, in the past such an evaluation task was often made difficult by the lack of accessible sources to gather the necessary information. With the rise of the internet, a large number of online platforms specialized on specific complex goods, e.g. used cars, makes it possible to easily cross-compare similar offers for close substitute objects. However, recent research (see, e.g., Lacetera *et al.* (2012)) documents that consumers fail to efficiently process the available relevant information but instead rely on heuristic evaluation rules.

Based on detailed field data on used car offers from the German website *mobile.de*, one of Europe's largest online vehicle marketplaces, we add to this evidence. We show that there are clear threshold effects on prices[1] at year changes in the date of first registration[2], i.e., a car's age. All else equal, the price differential between

[1] Our sample of used cars does not originate from an auction market, and, rather than on actual sales prices, our analysis is based on the asking prices stated by the individual sellers, which may be subject to negotiation once an interested buyer has been found. However, we have strong reasons to believe that the posted price is a sensible proxy for the final price in this market. First, *mobile.de* offers the seller an option to declare the stated price either as "fixed" or as "negotiable", and a substantial fraction of the sellers opts for the former rather than the latter. Second, with several thousand offers for each model series the market for used cars is highly competitive. Moreover, the cars within each of our subsamples can be regarded as close substitutes. Under the presumption, that the stated sales price reflects the willingness to accept of the respective seller, according to Hanemann (1991) and Shogren *et al.* (1994) in such an environment an endowment effect, i.e. a divergence of willingness to pay and willingness to accept, is unlikely to persist. Moreover, the services of *mobile.de* are widely used by professional car dealers who purchase cars for resale rather than use, where according to Kahneman *et al.* (1991) the endowment effect does not apply. As it turns out, the majority of offers in our sample is indeed made by commercial rather than private sellers. Hence, it stands to reason that the stated prices are closely related to the final prices. Finally, since advertising a car is costly, it seems plausible that the sellers exert considerable effort to elicit a reasonable price, at which prospective buyers are indeed willing to buy.In line with this argument, Englmaier and Schmöller (2009a) document that the sellers' reserve prices in a similar, but distinct, online-auctions market are similarly determined as the sales prices, i.e. from an evaluation of the individual attributes. Our intuition is that the same also applies to this context. For simplicity, in the following we use the term "price" to refer to the stated prices in our data.

[2] In Germany, every car has a legally mandated, official documentation record. Hence, the date of first registration, i.e. a car's birthdate for being on the road, is verifiable hard information.

two cars, where one was first registered in January and the other in December of the previous year, is dramatically larger than that between two cars first registered in any two subsequent months of the same year, respectively. Stated differently, we find an amplified adjustment in the prices for otherwise identical cars to be located *across* different registration years, or "vintages", where the impact of a marginal month of age is *up to four times larger* relative to that *within* the same vintage.

Moreover, replicating the findings in Lacetera *et al.* (2012), we find discontinuous price changes around 10,000km marks of odometer readings. For example, cars with odometer values between 59,500 and 59,999km are offered only slightly cheaper than cars with odometer readings between 59,000 and 59,499 but the price drop to the 60,000-60,499 bin is substantially bigger.

These discontinuous drops in value are evident in simple graphs of the raw data and are systematically confirmed by regression analyses, controlling for a host of observable car characteristics, including the exact mileage and age.

The discontinuities regarding to the odometer readings can be, as suggested by Lacetera *et al.* (2012), reconciled with a model of left-digit bias in information processing. However, this does not explain the discontinuities between a car fist registered in December (e.g., 12/2004) and January (e.g., 01/2005). We propose a model of (rational or heuristic) search costs, based on the design of the search interface where agents are led to search within pre-specified intervalls related to prominent marks.

The results in Englmaier and Schmöller (2009b) complement this explanation approach. In a different online market with similar features, Englmaier and Schmöller (2009b) can exploit an exogenous change in the search interface, amounting to a substantial reduction in search costs and find that the size of previously existing price discontinuities is substantially reduced.[3]

Limited attention has also been documented for other purchase decisions in other markets. For instance, Lee and Malmendier (2007) analyze individual bidding behavior in auctions on eBay and find that people tend to anchor on an irrelevant outside retail price for a board game, if the seller chose to state that price in the description of the product details. At the same time, many of the winning bids exceed a more relevant outside option, the so called "buy-it-now" price, which is an ex-ante fixed strike price set by the seller as an alternative to the auction

Moreover, in Germany there is no concept of "model year". I.e., there is no such thing as an 1998 BMW but there exists a certain BMW series that is produced without major changes for an extended period of several years.

[3]The paramount role of information provision in online markets is underlined by Lewis (2011). Tadelis and Zettelmeyer (2011) document it for the used car markt also studied by Lacetera *et al.* (2012).

process. Analyzing stock market data, Gilbert *et al.* (2008) provide evidence that investors with limited attention have an incentive to focus on summary statistics rather than individual pieces of information. They analyze the market response to the U.S. Leading Economic Index (LEI), a macroeconomic release that is purely a summary statistic, and show that the LEI announcement has an impact on aggregate stock returns, return volatility, and trading volume. We add to these findings by demonstrating that inattentiveness effects pertain for complex goods and large stake purchase decisions, even though the concerned piece of information is provided at arm's length within the relevant market environment.

The remainder of the paper is structured as follows. Section 2 describes the structure of the data and the relevant details of the sample selection. Section 3 presents our graphical analysis and the regression analysis. Section 4 presents a simple model to rationalize our results. Section 5 concludes and the Appendix collects all Tables and Figures.

2 Data Description

2.1 Institutional Background

For the purpose of this study, we collected detailed information on more than 80,000 cars offered during July and August 2009 on the online vehicle market platform *mobile.de*. Founded in 1996, *mobile.de* takes the role of an intermediator between supply and demand within a two-sided market. The company itself is not involved at any stage in the purchase or sale of a vehicle and a successful sale does not invoke any final value fees to *mobile.de*. It provides both a platform for sellers to place advertisements for new and used cars at a small cost and a free comprehensive search tool for prospective buyers to screen among the mass of on average about 1.3 million offers. According to the company's own statement, prospective buyers "*can limit search results by setting individual preferences and like this obtain customized offers with just a few clicks*", providing them "*. . . with an overview of the market and information about prices*".[4] The same is true for a seller who wants to evaluate his car before placing a sales advertisement.

`-- Include Figure 1 about here. --`

Figure 1 shows the interface a user is presented with upon entering *mobile.de*'s website. It displays a simple search form, which among other things allows to filter for makes, models, and a number of other basic details. A detailed search

[4]Source: http://cms.mobile.de/en/company/portrait_mobile.html

form, which can be directly reached by clicking the link to the lower left, provides a large additional set of filter options. Note however, that the drop down selector for the date of first registration only allows to filter for the vintage, i.e. the year of first registration. Neither in the simple nor the expanded form it is possible to adjust the search inquiry for the precise month of the first registration (FR).

The search returns a list of all vehicles matching the chosen filters. Per default they are sorted by price, where an abstract of their main features is displayed as shown in Figure 2. This preview explicitly states the precise date of first registration (e.g. "FR 01/2000") and additionally provides valuable information on the price, mileage, color, and power of the car, to name only a few. It is also possible to remember a specific car for later access ("Park vehicle"), which allows the user to directly compare the latter to other remembered cars.

-- Include Figure 2 about here. --

A typical profile page of an offered car, which is accessed from the search results list by clicking on the model name at the top of the respective entry, is depicted in Figure 3.

-- Include Figure 3 about here. --

For each car, a seller has to specify a preselected set of features and attributes, where most of the respective values are chosen from a drop down menu during the preparation of the advertisement. Conveniently, this data is thus standardized and ensures a sufficient degree of comparability across individual observations. We therefore focus on these standard attributes in our data, which in addition to the stated price and the date of FR include various extras and also some information on the sellers (see Tables 2, 3, 4, and 5 below).

2.2 Sample selection

Our data includes details on the most widespread car models from four leading German makes[5], all ranked among the top seven of Germany's vehicle population according to the Kraftfahrt-Bundesamt (KBA).[6] More specifically, we collected information on 29,097 Volkswagen (VW) Golf (KBA-rank 1), 14,693 Opel Astra (KBA-rank 2), 25,582 BMW 3 (KBA-rank 4), and 17,901 Audi A4 (KBA-rank

[5]We had to limit our search and we chose to focus on the most widespread brand of each of the four biggest car producers in Germany to avoid potential specifities in sub-markets related to certain brands.

[6]Source: http://www.kba.de.

7), all advertised as accident-free and with their FR-dates between 01/2000 and 12/2008.[7] We focus on this subsample for two main reasons. First, a high stock is a good indicator for a considerable volume of used car offers for a specific model, which ensures a sufficiently large number of observations. Second, we consider models from different makes to achieve a broad diversification within our identification strategy.[8]

-- Include Table 1 about here. --

Since the introduction of a new series within a particular car model affects the sales prices substantially, we can only retrieve meaningful estimates of the influential attributes if we accurately control for potential model revisions. Clearly, this requires detailed knowledge of the exact dates of the respective market launches. Conveniently, for the four different models considered in our sample, this information is readily available. In particular, we identify the respective estimation windows for each model according to the information provided through the manufacturers' websites, the Schwacke-List (http://schwacke.de), and the Deutsche Automobil Treuhand (http://www.dat.de).[9][10]

Naturally, due to different variants offered within a model series, e.g. limousine, estate car, or compact car, the latter are not perfect substitutes. To account for such within-series variation, we distinguish between five- and three door versions, add a large set of main attributes as controls, and exclude convertibles from the sample. In this way, we capture a substantial share of the variation in the price within a series and are thus able to obtain precise estimates of the influential factors.

[7] KBA-ranks not reported were taken by other models of VW (Passat, Polo) and Opel (Corsa).

[8] We do not consider cars that were first registered before January 1998, since their values are very low. Moreover, we thereby avoid a potential "left-digit" effect with respect to the registration year: It has been documented that some individuals tend to process numerical information in a way that the first digits are treated as more valuable information, i.e. are perceived to contain more significant information than later digits (see e.g. Brenner and Brenner, 1982; Bhattacharya *et al.*, 2008; Lacetera *et al.*, 2012).

[9] The latter are commercial service providers who offer benchmark evaluations for all kind of cars at a small cost. In fact, they allow to account for the precise date of first registration in an individual evaluation of a car, which makes the discontinuities we are able to document in our data even more puzzling.

[10] Since all models in consideration experienced at least one upgrade or change of series between 2000 and 2008, we define the estimation periods accordingly and are thus able to conduct our analysis for eight different subsamples, as shown in Table 1. These disaggregate analyses do confirm our finding from the pooled regressions reported in the paper. To save on space, we do not report the results of these separate analyses here. Depending on their extent, these upgrades, or "face-lifts", can invoke similar price effects as a change of series. In the estimation we therefore treat the information on a face-lift similar to the introduction of a new production line.

-- Include Table 2 about here. --

-- Include Table 3 about here. --

-- Include Table 4 about here. --

-- Include Table 5 about here. --

2.3 Data description

Tables 2, 3, 4, and 5 provide an overview of the collected details and show the corresponding summary statistics. Since a complete description of all subsamples would go beyond the scope of the paper, throughout the next few paragraphs we representatively focus on the samples of VW Golf series IV and V.

In general, the value of an individual car from a specific model series depends on numerous factors. Among others, this includes its age, its odometer reading, the power and fuel-type of its engine, and the different extras it is equipped with, e.g. an automatic gearbox, a sun-roof, a seat-heating, or a cruise control. Along with the stated prices and the month and year of first registration, we therefore collected a large number of features for each of the cars to control directly for quality differences. To measure their impact on the price of the car, we assign a dummy variable to each of the observed extras in our analysis. For instance, if a offered car has a sun-roof, the dummy variable *sun_roof* takes value 1 and 0 otherwise.[11] Similarly, we also add a dummy for both the door-count and the fuel type.[12]

The information on the month and year of the first registration is stored in the variables $fr\_month \in [1, 12]$ and $fr\_year \in [2000, 2008]$, respectively. For our empirical analysis, we combine the latter to construct the measure $totalage \in [1, 108]$, which displays the precise age of a car in units of months:

$$totalage \equiv 12 \cdot (2008 - fr\_year) + (13 - fr\_month),$$

where the normalization is such that a car's age is measured relative to the most recent FR-date included within our dataset, i.e. 12/2008, which corresponds to the minimum age of 1 month.

Returning to Table 2, observe that the average series IV (V) Golf has an age of 87 (30) months, exhibits an odometer reading of 115,978 km (55,343 km) and

[11] In the following, we use italics to denote the variable name in our data corresponding to an attribute.

[12] For air conditioning, airbags, and electric window lifters we find almost no variation in the data. Since by now these features are included in the basic configuration of most cars, we omit them from the analysis.

is offered at a price of EUR6,711 (EUR13,605), as shown in Panel A (B). In both series, most of the cars have five doors, a manual gearbox, and a metallic paint. However, the frequency of diesel cars, seat heatings, all-wheel drives, and cruise controls is considerably higher for the newer series V than for series IV. Also note that a large majority of offers originates from professional car dealers, as indicated by the dummy *private_seller* being equal to zero.

As we would expect, a correlation analysis for *price* yields a strong negative correlation coefficient with *totalage* ($\rho = -0.85$) and with *mileage* ($\rho = -0.78$). Conversely, *power* ($\rho = 0.45$), *diesel* ($\rho = 0.11$), *five-door* ($\rho = 0.16$), and all of the considered extras are significantly positively related to the price of a car.[13]

While not listed in Table 2, another important determinant for the price of a car is its color. We therefore additionally include a set of color-dummies to control for their impact on price, where the effects are measured relative to *black*. We find that the prices are indeed somewhat responsive to different colors. For the sake of clarity, however, in the discussion below the respective coefficients for the color-dummies are not reported, but are available from the authors upon request.

-- Include Figure 4 about here. --

-- Include Figure 5 about here. --

-- Include Figure 7 about here. --

Next, look again at our full sample and consider the age distribution of the cars, which are depicted in Figure 4. We find some fluctuation across registration months but conclude that our sample contains a sufficient number of observations for each FR-date in the estimation period. The highest frequency of offers is observed for relatively new cars, i.e. around an age of 7 to 15 months relative to 12/2008, which is not surprising given the high number of professional car dealers that is active in this market segment. In contrast, prices are approximately normally distributed (Figure 5). Finally, mileage is pretty much evenly distributed with a plateau below 30,000km and another fairly stable but substantially lower plateau above 30,000km (Figure 7).

[13] Among the explanatory variables, we find that *totalage* and *mileage* co-move at a degree of $\rho = 0.77$. While in general collinearity among the explanatory variables can be problematic, our sample size is sufficiently large to produce precise parameter estimates.

3 Empirical Analysis

3.1 Graphical Analysis

Age Discontinuities We begin the empirical analysis by simply plotting the raw price data as a function of car age. In Figures 8ff, each dot shows the average sale price for all cars first registered in a given month of a given year starting December 2008 and counting backwards until January 1998.

-- Include Figure 8 about here. --

As one would expect, average prices decrease with increasing age. Within each year, monthly average prices decline almost linearly, but there are marked discontinuities between years (Figure 8). These patterns are very systematic (and substantial) for all cars at least 2 years old (Figure 9).[14]

-- Include Figure 9 about here. --

With no other explanation for the importance of the year thresholds, these results suggest a role for limited attention in this market.

Mileage Discontinuities Next we plot the raw price data as a function of car mileage.In Figures 11ff, each dot shows the average sale price for cars in a 500-km mileage bin, starting at 1,000 km. There is a dot for the average price of cars with 1,000 through 1,499 km, then a dot for cars with 1,500 to 1,999 km, and so on. The vertical lines in the graph indicate each 10,000-km mark.

As one would expect, average prices decrease with increasing mileage. Within each 10,000-mile band, average prices decline quite smoothly. Focussing on the thickest market segment of cars with medium mileage between 20,000 - 60,000 km, there are clear and sizeable discontinuities in average prices at all 10,000-km marks (Figure 12).

-- Include Figure 12 about here. --

These price discontinuities are also present for high and low mileage cars, but the pattern is less systematic (13 and 14).

With no other explanation for the importance of 10,000-km thresholds, these results, too, strongly suggest a role for limited attention in this market.

[14]The "youngest" cars are a specific sub-sample. Namely so called "Jahreswagen", i.e. cars given for a year to employees as part of their compensation package. For these cars, there is also a marked discontinuity upon the year change, but in the "wrong" direction. (Figure 10).

3.2 Regression Analysis

The graphical analysis in the preceding subsection established the existence of systematic price discontinuities at year changes for the month of first registration and around 10,000km thresholds of odometer readings. To augment this approach, we turn now to regression analysis to establish numerical estimates of these price discontinuities. We implement regression discontinuity designs - see Lee and Lemieux (2010) for an overview of this literature - where the dependent variable in our regressions is the price for the cars stated on the website.

Age Discontinuities To establish the effect of crossing a year threshold, we control for the actual age - by means of a fifth-order polynomial[15] - , the mileage and other relevant characteristics of the particular car being sold like make, model, special equipment, etc..

The regression also includes a series of indicator variables for whether the car has crossed a given year threshold. The coefficients of these indicator variables can be interpreted as the discontinuous changes in price (all else constant) that occur as cars cross a particular year threshold. Hence, the specification allows us to estimate the price discontinuities separately at each year threshold.

```
-- Include Table 6 about here. --
```

Table 6 presents the regression results for the above described specification. The first column controls only for the full set of indicator variables for whether the car has crossed a given year threshold and provides estimates of the price discontinuities before any further controls on observables. Given our sizeable N, the coefficients are highly statistically significant and all of the coefficient estimates are negative, which is consistent with our inattention story.

Columns 2 through 4 in the table add a fifth-order age polynomial and increasingly restrictive fixed effects to the model. Column 2 adds a fifth-order age polynomial which actually affects significance, size and even sign of the coefficients. Sill, the results indicate that the threshold indicators are a reasonable approximation for the continuous polynomial. Adding controls for car features in Column 3 and car model fixed effects in Column 4 increases R^2. With the controlled estimates, most of the coefficients are highly significant, rather big (on average EUR 300), and negative.

[15] The specific functional form was chosen based on the Akaike Information Criterion test. Our results are robust to the specific order of the polynomial.

Though the pattern is not perfect, by and large we conclude that there are sizeable and systematic negative price discontinuities upon passing a year threshold even after controlling for the exact age and a host of observable characteristics.

Mileage Discontinuities To establish the effect of crossing a mileage threshold, we control for the actual mileage - by means of a third-order polynomial[16] - , the age and other relevant characteristics of the particular car being sold like make, model, special equipment, etc..

The regression also includes a series of indicator variables for whether the car has crossed a given 10,000km mileage threshold. The coefficients of these indicator variables can be interpreted as the discontinuous changes in price (all else constant) that occur as cars cross a particular 10,000km mileage threshold. Hence, the specification allows us to estimate the price discontinuities separately at each mileage threshold.

-- Include Table 7 about here. --

Table 7 presents the regression results for the above described specification. The first column controls only for the full set of indicator variables for whether the car has crossed a given 10,000km mileage threshold and provides estimates of the price discontinuities before any further controls on observables. Given our sizeable N, the coefficients are generally statistically significant and almost all of the coefficient estimates are sizeable and negative, which is consistent with our inattention story.

Columns 2 through 4 in the table add a third-order mileage polynomial and increasingly restrictive fixed effects to the model. Column 2 adds third-order mileage polynomial which again affects significance, size, and even sign of the coefficients. Still, the results indicate that the threshold indicators are a reasonable approximation for the continuous polynomial. Adding controls for car features in Column 3 and car model fixed effects in Column 4 increases R^2. With the controlled estimates, the coefficients for the lower 10,000km mileage thresholds (up to 40,000km) are negative, large (EUR 400 - 1200), and highly significant. Moreover, all higher mileage thresholds are also negative, though not significantly so. We conclude that there are sizeable and systematic negative price discontinuities upon passing a 10,000km mileage threshold even after controlling for the exact mileage and a host of observable characteristics.

[16]The specific functional form was chosen based on the Akaike Information Criterion test. Our results are robust to the specific order of the polynomial.

4 A simple model of costly search

In light of a rather sophisticated behavior as displayed with respect to other attributes, the main finding we are able to document in our data seems even more puzzling. If people are careful enough to check out numerous details of the attribute vector of a car, why do they systematically pay too little attention to the valuable information conveyed through the month of first registration? Though they do not disregard the impact of precise age as indicated by the evidential continuous decline *within* a vintage, they fail to recognize the connection to subsequent or previous vintages. Our intuition is that individuals evaluate cars relative to the average car from the same vintage, while the more relevant and informative peer group consists of cars of close-by actual age, irrespectively of the vintage the latter belong to.

To illustrate what we have in mind, consider an agent who has to evaluate a car with given attributes first registered in 12/2006. All else equal, to elicit how much to bid for this car, she should look up and compare the prices for a car of a similar total age, say, roughly from 3 months younger (09/2006) to three months older (03/2007). To get this information the agent has to screen a large number of cars on the market platform to find enough falling into that age-range, thus involving a time consuming and thus costly search. This implies that one possible explanation for the inattentiveness effect lies within the design of the user interface of the search engine, which is used to screen the market: It may be physical search costs that prevent an agent from efficient information aggregation.

Consider a risk-neutral agent j who wants to evaluate a particular car $i = (y_i, m_i, X_i)$, where $y_i \in \{2000, ..., 2008\}$ denotes its vintage, $m_i \in \{0, ..., 12\}$ the month of first registration, and X_i all other attributes of the car, respectively. Normalize by $a_i = 12 \cdot (2008 - y_i) + (13 - m_i)$ the total age in days. For given values of $X_i = \overline{X}$, individual j's value estimate for car i in dependence of its age attribute is described by the function $E_j[v_i] : (y_i, m_i) \to \mathbb{R}^+$.[17] More specifically, let

$$E_j[v_i] := (1 - \pi(c_j)) \cdot \overline{v}^j_y + \pi(c_j) \cdot v^j_{a_i},$$

where $\overline{v}^j_y$ is the value of an average car in age-group y, and $v^j_{a_i}$ denotes her precise value of car i. For simplicity, assume that $\overline{v}^j_y$ is commonly available free of cost. Her value estimate is a convex combination of the average value and her true value, where the relative weight $\pi(c_j)$ is a function of her search costs c_j. By screening the market for otherwise identical cars within an age-range around a_i, she can learn their values and thus increase the weight $\pi(\cdot)$ on her true value for

[17] To simplify the notation we suppress $\overline{X}$ in the expressions.

car i and thereby obtains a more precise estimate.[18] Generally, the intensity of this search will depend on how costly, or time consuming, it is to find appropriate cars in the respective age interval. Formally, assume that the convex weighting function $\pi(c_j)$ has the following properties:

$$\begin{aligned} \lim_{c_j \to \infty} \pi(c_j) &= 0 \\ \lim_{c_j \to 0} \pi(c_j) &= 1 \\ \pi'(c_j) &< 0 \quad \forall \ c_j \in R_0^+. \end{aligned}$$

First, consider that the search costs are sufficiently large such that $\pi(c_j) = 0$. Then manager j's valuation will reflect the average value $\overline{v}_y^j$. Second, for a given c_j suppose that $\pi(c_j) < 1$. If $v_{a_i}^j > \overline{v}_y^j$, the agent values the car too low, though her true value for the car would be higher than her estimate. Conversely, if $v_{a_i}^j < \overline{v}_y^j$, she will pay a price above her true valuation for car i. While the former case is unproblematic, in the latter the agent with the least precise estimate will determine the final price. Third, in the absence of search costs, agent j will fully learn her precise value, i.e. $\pi(0) = 1$. These cases are illustrated in Figure 15.

-- Include Figure 15 about here. --

For lower search cost $c_j' < c_j$, the estimates of any individual agent j should become more accurate in the sense that they become closer to her precise value $v_{a_i}^j$ since $\pi(c_j) < \pi(c_j')$. Our data are thus consistent with a model that includes intermediate search costs, $0 < \pi < 1$, which suffice to cause discontinuities between two consecutive vintages. Obviously, a similar reasoning can be applied to explain the price drops around 10,000km odometer marks.

5 Discussion and Conclusion

We examine empirically to what extent the stated prices for used cars reflect available and relevant information. Based on detailed field data on used car offers from the online vehicle market platform *mobile.de*, we find strong evidence for biased information processing. Despite the large monetary stakes involved, our findings suggest that people in this market systematically fail to efficiently aggregate the

[18] The underlying rationale may be best explained by assuming that for any c_j, the agent solves an optimal search problem, which determines the number of cars she optimally screens. In turn, this implicitly determines the extent to which she learns $v_{a_i}^j$.

information provided on specific attributes of the items on sale. In particular, although the precise date of first registration is clearly stated, the pattern of observed prices exhibits sizeable discontinuities, indicating that a substantial fraction of the value adjustment due to the age of a car is located where the FR-year changes. As a consequence, across two consecutive vintages the price differential for cars with otherwise close-by registration dates is significantly larger than rationally justified, given that they only marginally differ in their precise age. This finding proves robust. Moreover, we are able to replicate the findings from Lacetera *et al.* (2012) and find discontinuous drops in prices at 10,000km odometer thresholds.

The fact that we are able to provide suggestive evidence for a systematic friction in an otherwise highly competitive market, where in addition individual choices are conceivably subject to profound deliberations, naturally raises two closely related questions. First, what are the driving forces behind this effect? And second, what are the economic consequences of this finding?

Regarding the first question: While the latter finding, as suggested by Lacetera *et al.* (2012), is consistent with a left-digit bias in the processing of numerical information, the first finding cannot be explained by this. We suggest a model of (rational or heuristic) search costs that is capable of explaining bot price patterns.

Regarding the second question, these price discontinuities might entail that from the perspective of rational buyers a substantial fraction of cars will be overpriced, potentially leading to too little trade. Or, from the perspective of rational sellers, cars from some segments will appear underpriced, potentially leading to too little trade from the supply side.

Several extensions to this research suggest themselves. One potential source for this effect may be linked to the design of the filter mechanism, which the people can use to screen and cross compare different offers. Due to the fact that it is not possible to directly filter for the FR-month on platforms like *mobile.de*, it may be tempting to perceive this information as unimportant and to overly focus one's attention on the more salient FR-year. It would therefore be interesting to see whether the size of the discontinuities is affected by including this feature in the filter mechanism. In a different online market with similar features but for small stakes items, Englmaier and Schmöller (2009b) can exploit an exogenous change in the search interface, amounting to a substantial reduction in search costs and find that the size of previously existing price discontinuities is substantially reduced.

In his seminal contribution to information economics, Akerlof (1970) employs the information asymmetries between buyers and sellers of used cars as his prime example to illustrate the famous "lemons-problem". Although adverse selection

due to asymmetric information with respect to unobservables is undeniably still a major problem within this market, our findings suggest that inefficiencies may also arise with respect to observable characteristics. People seem to be inattentive to subtle, but nevertheless valuable details of the *available* information.

References

Akerlof, G. A. (1970). The market for "lemons": Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics*, **Vol. 84(3)**, pp. 488–500. 13

Bhattacharya, U., Holden, C. W. and Jacobsen, S. E. (2008). Penny wise, dollar foolish: The left-digit effect in security trading. *Working Paper*. 5

Brenner, G. and Brenner, R. (1982). Memory and markets, or why are you paying $2.99 for a widget? *The Journal of Business*, **Vol. 55**, pp. 147–158. 5

Englmaier, F. and Schmöller, A. (2009a). Determinants and effects of reserve prices in hattrick auctions. *Working Paper*. 1

— and Schmöller, A. (2009b). Does bidding for complex goods reflect all relevant information? field evidence from online gaming. *Working Paper*. 2, 13

Gilbert, T., Kogan, S., Lochstoer, L. A. and Ozyildirim, A. (2008). Investor inattention and the market impact of summary statistics. *Working Paper*. 3

Hanemann, W. M. (1991). Willingness to pay and willingness to accept: How much can they differ? *American Economic Review*, **Vol. 81** (3), pp. 635–647. 1

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

Lacetera, N., Pope, D. G. and Sydnor, H. (2012). Heuristic thinking and limited attention in the car market. *American Economic Review*, **102(5)**, 2206–2236. 1, 2, 5, 13

Lee, D. S. and Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature*, **48(2)**, 281–355. 9

Lee, Y. H. and Malmendier, U. (2007). The bidder's curse. *NBER Working Paper No. 13699*. 2

Lewis, G. (2011). Asymmetric information, adverse selection and online disclosure: The case of ebay motors. *American Economic Review*, **101 (4)**, 1535–1546. 2

Shogren, J. F., Shin, S. Y., Hayes, D. J. and Kliebenstein, J. B. (1994). Resolving differences in willingness to pay and willingness to accept. *The American Economic Review*, **Vol. 84** (1), pp. 255–270. 1

Tadelis, S. and Zettelmeyer, F. (2011). Information disclosure as a matching mechanism: Theory and evidence from a field experiment. *Working Paper*. 2

A Figures

Figure 1. *Website www.mobile.de*

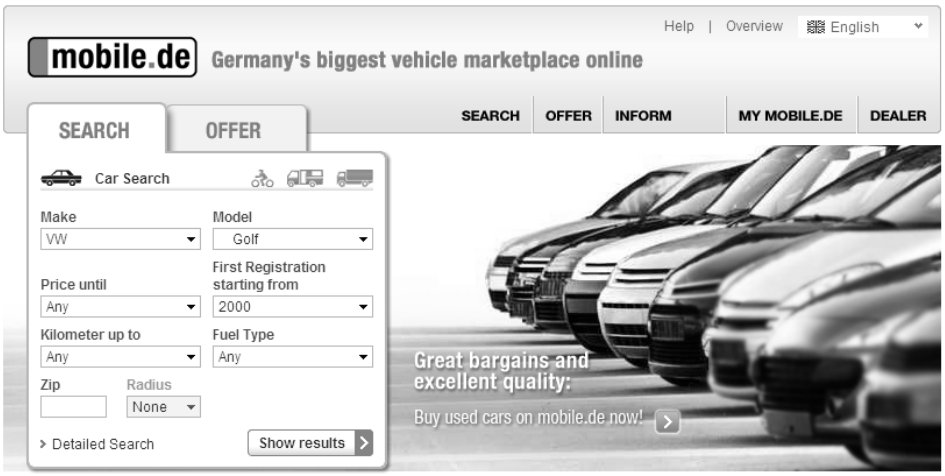


Figure 2. *www.mobile.de - Search Results*

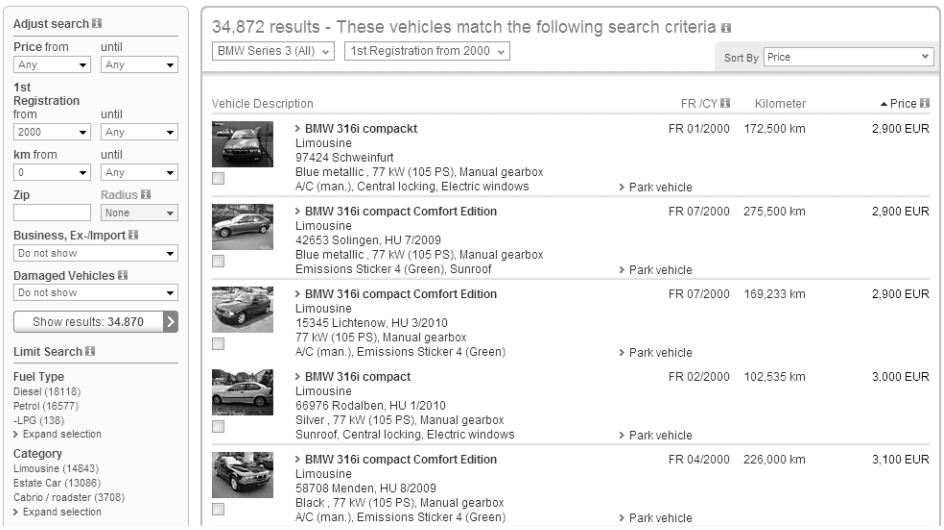


Figure 3. *www.mobile.de - Car Details*

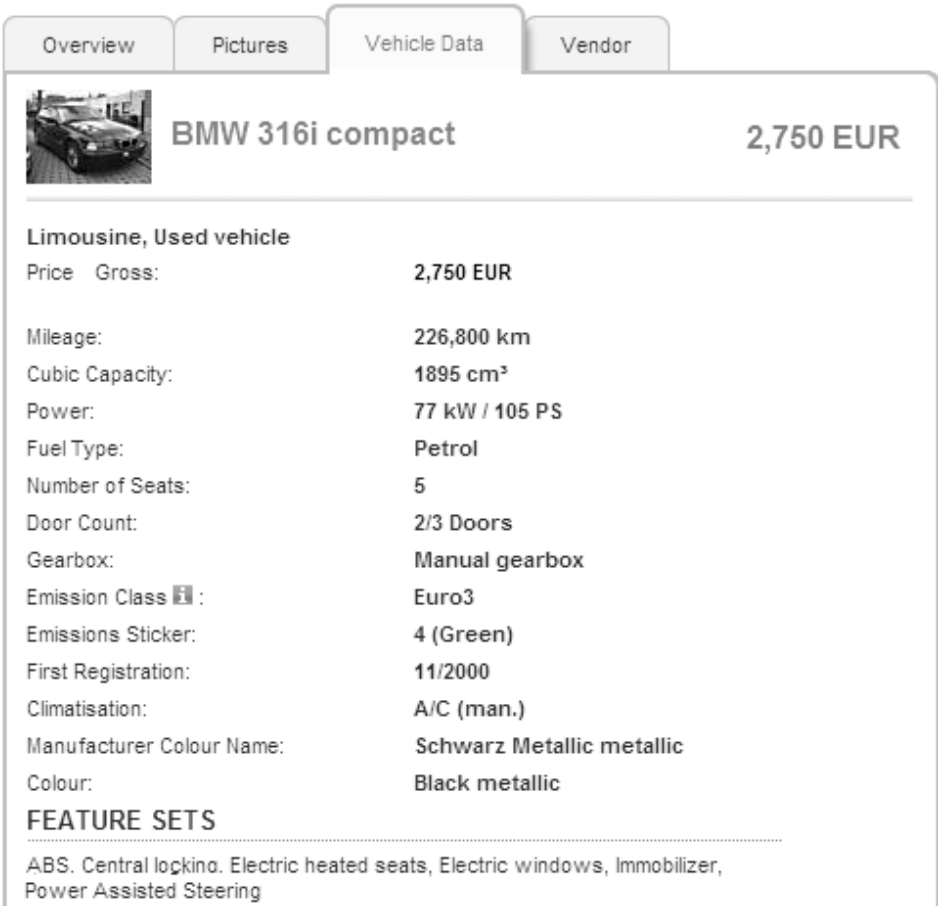


Figure 4. *Distribution of Car Age*

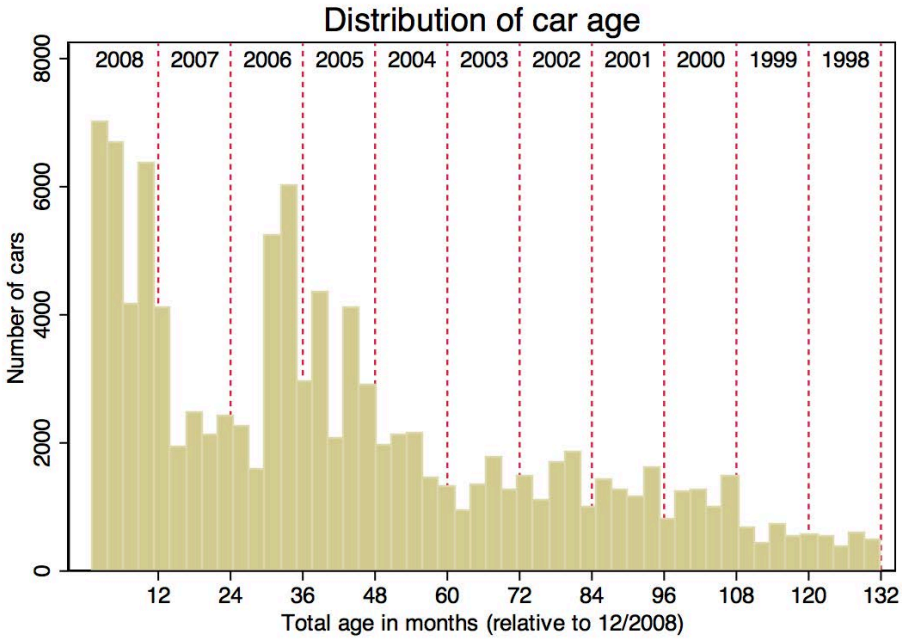


Figure 5. *Distribution of Prices (all cars)*

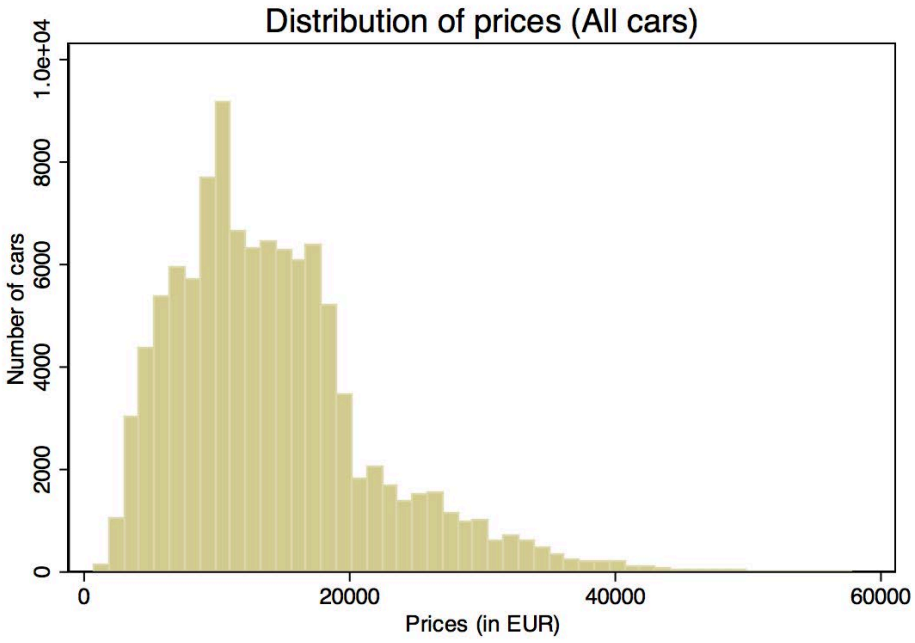


Figure 6. *Distribution of Prices (mid aged cars only)*

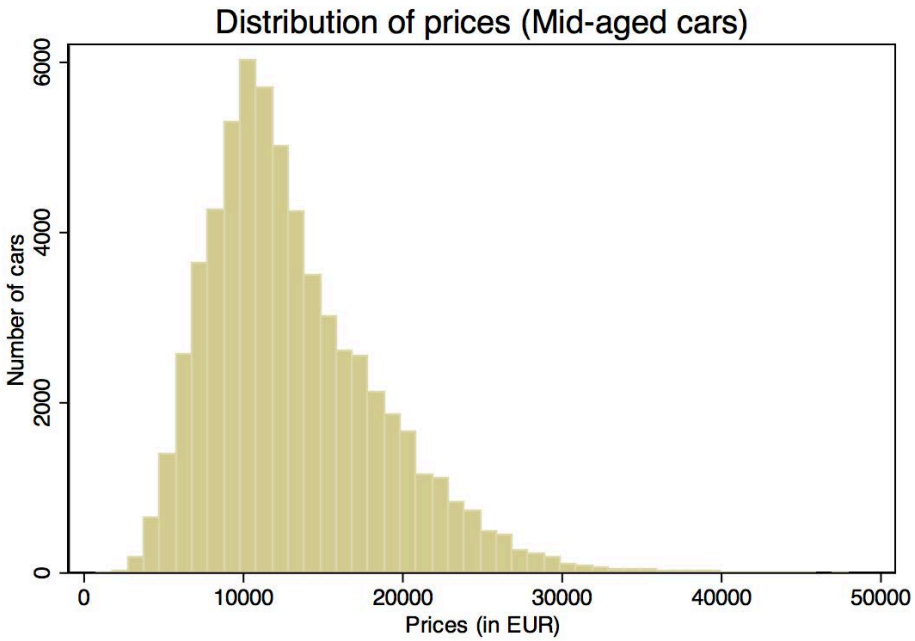


Figure 7. *Distribution of Mileage*

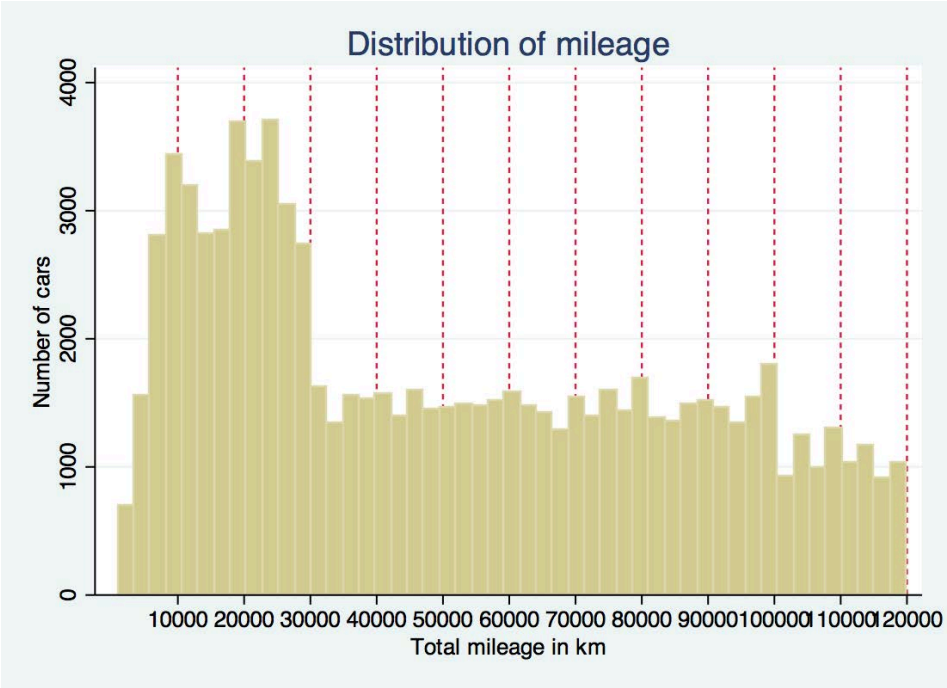


Figure 8. *Average Car Prices by Age (all cars)*

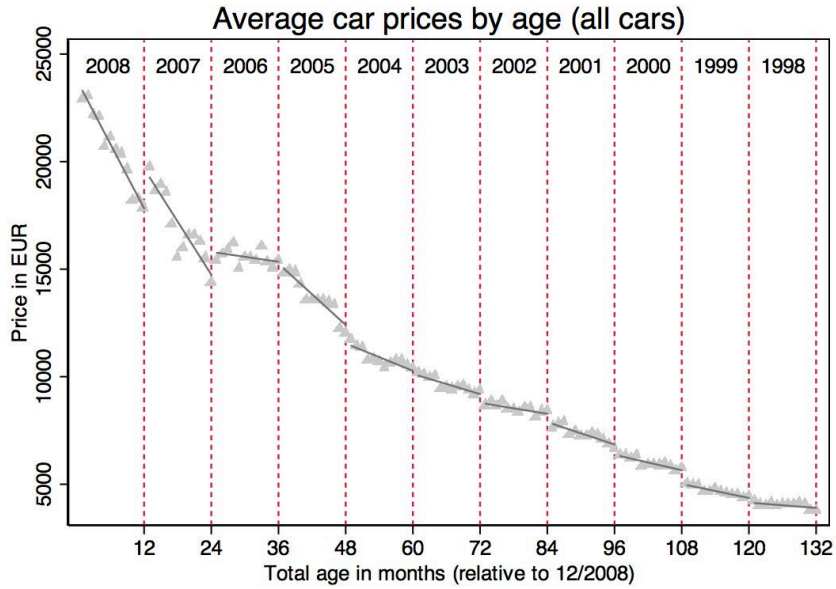


Figure 9. *Average Car Prices by Age (w/o new cars)*

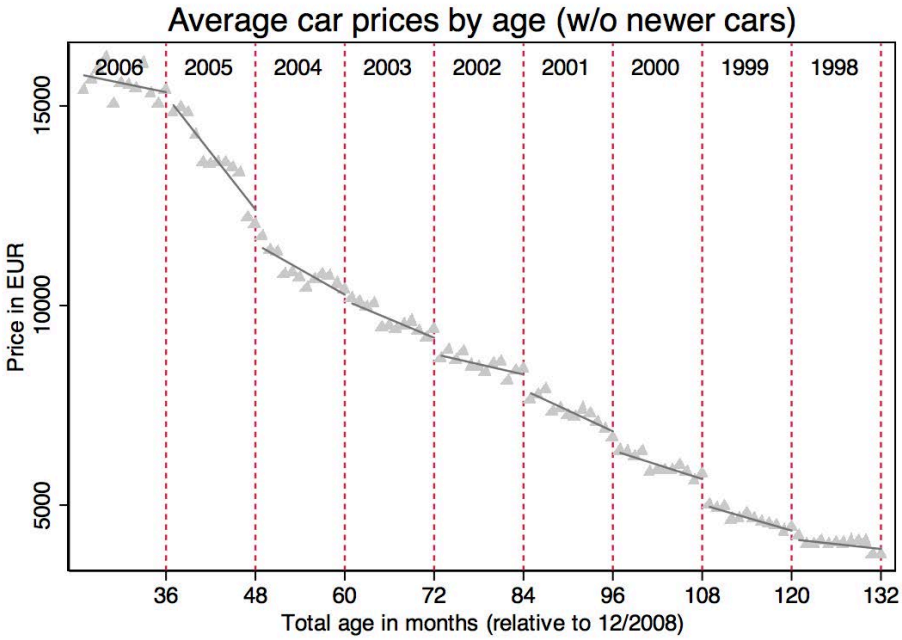


Figure 10. *Average Car Prices by Age (new cars only)*

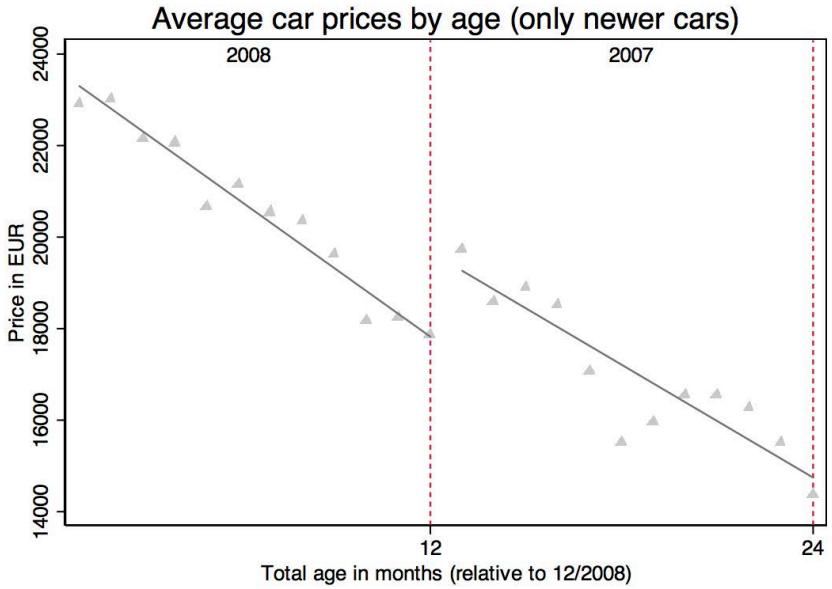


Figure 11. *Average (per 500Km bins) Car Prices by Mileage (all cars)*

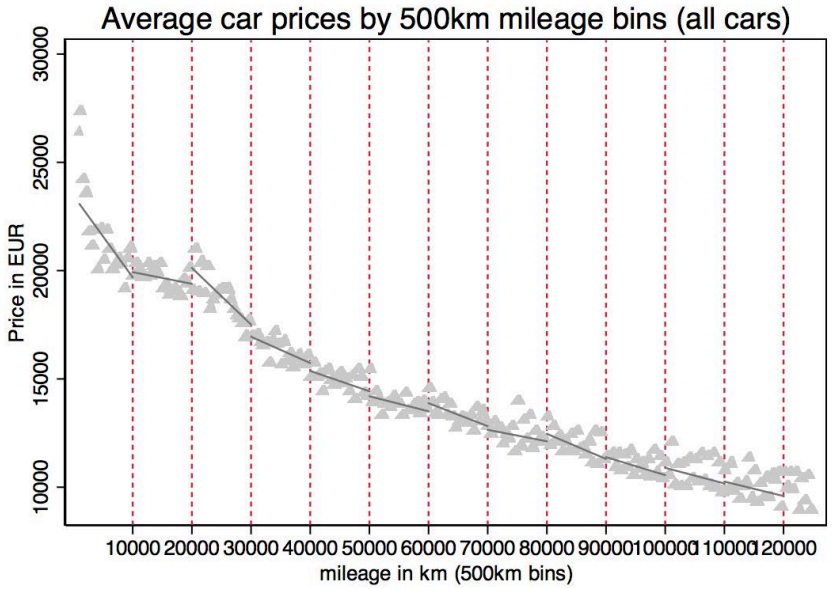


Figure 12. *Average (per 500Km bins) Car Prices by Mileage (medium mileage cars)*

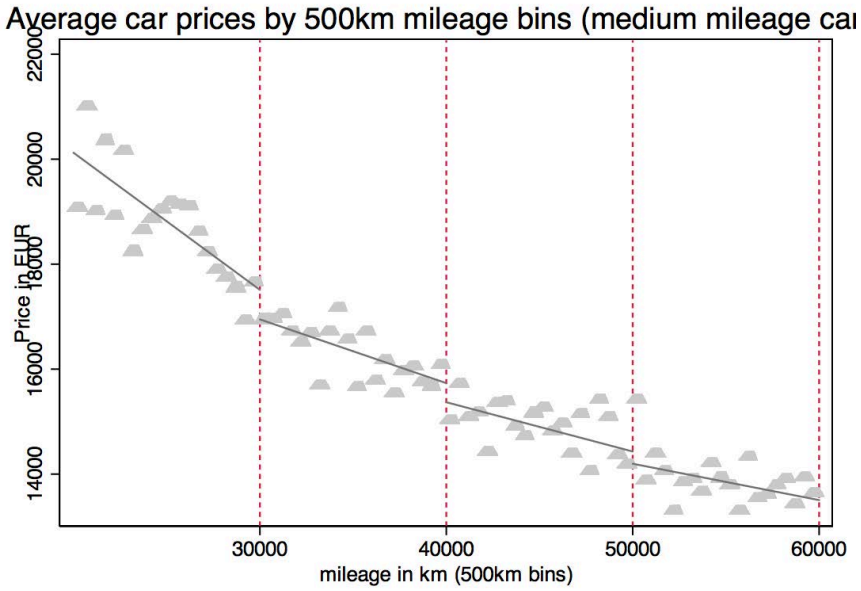


Figure 13. *Average (per 500Km bins) Car Prices by Mileage (high mileage cars)*

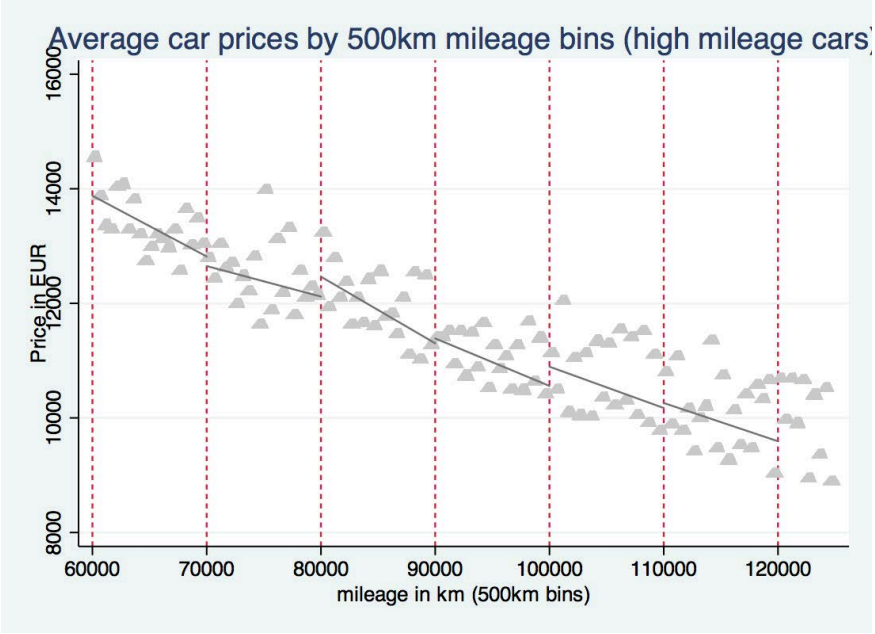


Figure 14. *Average (per 500Km bins) Car Prices by Mileage (low mileage cars)*

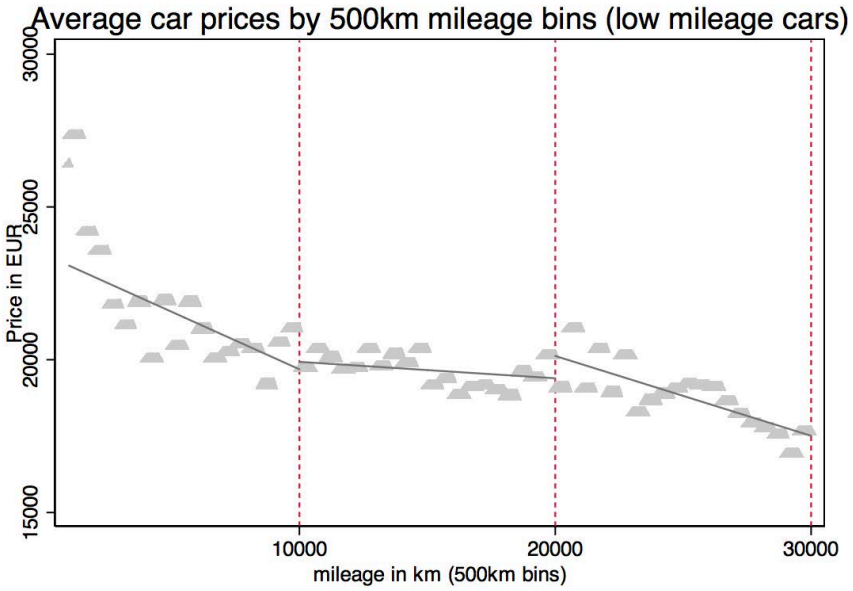
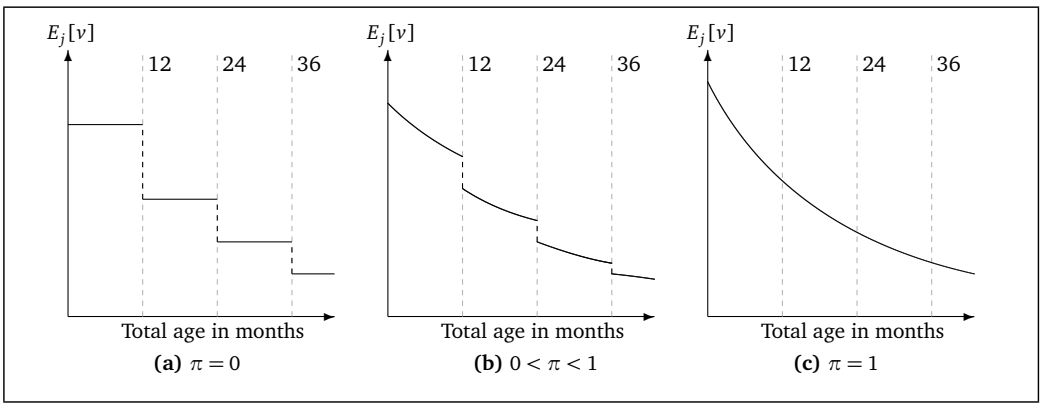


Figure 15. *Expected Valuation in Dependence of Search Costs*



B Tables

Table 1. *Models Series and Estimation Periods*

Make & Model	Name of Series		Production period	Estimation period
VW Golf	IV		10/1997 - 09/2003	01/2000 - 09/2003
	V		10/2003 - 07/2008	01/2004 - 06/2008
BMW 3	E46		04/1998 - 11/2004	01/2000 - 11/2004
	E90	(limousine)	12/2004 - 09/2008*	09/2005 - 08/2008
		(estate)	06/2005 - 09/2008*	
Audi A4	B6	(limousine)	10/2000 - 11/2004	10/2001 - 10/2004
		(estate)	09/2001 - 11/2004	
	B7	(limousine)	11/2004 - 11/2007	04/2005 - 09/2007
		(estate)	11/2004 - 03/2008	
Opel Astra	G		02/1998 - 01/2004	01/2000 - 12/2003
	H		02/2004 - 10/2007*	05/2004 - 10/2007

Notes: Entries with an asterisk indicate an upgrade of the current production series. If there were different introduction dates within a model series, we use the later date to determine the estimation period.

Table 2. *Summary Statistics - VW Golf*

	Panel A. Series IV (Est. Period: 01/2000 - 09/2003)				Panel B. Series V (Est. Period: 01/2004 - 06/2008)			
Variable	Obs.	Mean	Min	Max	Obs.	Mean	Min	Max
price [EUR]	6,034	6,711	2,499	17,990	15,247	13,605	4,900	27,979
mileage [km]	6,034	115,978	2,000	225,500	15,247	55,343	1,000	221321
power [kW]	6,034	73	50	213	15,247	81	50	243
fr_year	6,034	2001	2000	2003	15,247	2006	2004	2008
fr_month	6,034	6	1	12	15,247	6	1	12
totalage [months][a]	6,034	87	64	108	15,247	30	7	60
Dummies[b]	Value	Frequency	Percent	Cum.	Value	Frequency	Percent	Cum.
diesel	0	3,771	62.5	62.5	0	5,998	39.3	39.3
(0 = petrol, 1 = diesel)	1	2,263	37.5	100.0	1	9,249	60.7	100.0
five-door	0	1,656	27.4	27.4	0	2,921	19.2	19.2
(0 = three, 1 = five)	1	4,378	72.6	100.0	1	12,326	80.8	100.0
auto gearbox	0	5,509	91.3	91.3	0	13,364	87.7	87.7
(0 = manu., 1 = auto)	1	525	8.7	100.0	1	1,883	12.4	100.0
cruise control	0	5,222	86.5	86.5	0	7,895	51.8	51.8
(0 = no, 1 = yes)	1	812	13.5	100.0	1	7,352	48.2	100.0
seat heating	0	4,554	75.5	75.5	0	7,712	50.6	50.6
(0 = no, 1 = yes)	1	1,480	24.5	100.0	1	7,535	49.4	100.0
all-wheel drive	0	5,879	97.4	97.4	0	7,712	50.6	50.6
(0 = no, 1 = yes)	1	155	2.6	100.0	1	7,535	49.4	100.0
sun roof	0	5,106	84.6	84.6	0	13,634	89.4	89.4
(0 = no, 1 = yes)	1	928	15.4	100.0	1	1,613	10.6	100.0
leathertrim	0	5,817	96.4	96.4	0	14,560	95.5	95.5
(0 = no, 1 = yes)	1	217	3.6	100.0	1	687	4.5	100.0
metallic paint	0	1,589	26.3	26.3	0	3,258	21.4	21.4
(0 = no, 1 = yes)	1	4,445	73.7	100.0	1	11,989	78.6	100.0
private seller	0	4,419	73.2	73.2	0	13,821	90.7	90.7
(0 = no, 1 = yes)	1	1,615	26.8	100.0	1	1,426	9.4	100.0

Notes: [a] *totalage* $\equiv 12 \cdot$(2008-fr_year)$+$(13-fr_month) displays the age of a car in units of months and is normalized such that the minimum $1 \mathrel{\widehat{=}} 12/2008$ and the maximum $108 \mathrel{\widehat{=}} 01/2000$. For instance, the estimation period for VW Golf IV, i.e. 01/2000 to 09/2003, equals $totalage \in [64, 108]$. [b] Color dummies are not displayed to save on space.

Table 3. *Summary Statistics - BMW 3⋆*

	Panel A. Series E46 (Est. Period: 01/2000 - 11/2004)				Panel B. Series E90 (Est. Period: 09/2005 - 08/2008)			
Variable	Obs.	Mean	Min	Max	Obs.	Mean	Min	Max
price [EUR]	7,458	10,562	3,990	20,000	10,866	22,607	8,500	48,578
mileage [km]	7,458	113,130	6,163	227,000	10,866	58,220	1,000	226150
power [kW]	7,458	110	75	185	10,866	126	85	247
totalage [months]	7,458	77	50	108	10,866	26	5	40
Dummies	Value	Frequency	Percent	Cum.	Value	Frequency	Percent	Cum.
diesel	0	4,373	58.6	58.6	0	3,007	27.7	27.7
(0 = petrol, 1 = diesel)	1	3,085	41.4	100.0	1	7,859	72.3	100.0
auto gearbox	0	5,730	76.8	76.8	0	7,579	69.8	69.8
(0 = manu., 1 = auto)	1	1,728	23.2	100.0	1	3,287	30.3	100.0
cruise control	0	5,067	67.9	67.9	0	4,453	41.0	41.0
(0 = no, 1 = yes)	1	2,391	32.1	100.0	1	6,413	59.0	100.0
seat heating	0	3,793	50.9	50.9	0	3,309	30.5	30.5
(0 = no, 1 = yes)	1	3,665	49.1	100.0	1	7,557	69.6	100.0
all-wheel-drive	0	7,222	96.8	96.8	0	10,414	95.8	95.8
(0 = no, 1 = yes)	1	236	3.2	100.0	1	452	4.2	100.0
sun roof	0	4,735	63.5	63.5	0	6,848	63.0	63.0
(0 = no, 1 = yes)	1	2,723	36.5	100.0	1	4,018	37.0	100.0
leathertrim	0	5,359	71.9	71.9	0	7,684	70.7	70.7
(0 = no, 1 = yes)	1	2,099	28.1	100.0	1	3,182	29.3	100.0
metallic paint	0	1,091	14.6	14.6	0	2,957	27.2	27.2
(0 = no, 1 = yes)	1	6,367	85.4	100.0	1	7,909	72.8	100.0
private seller	0	5,544	74.3	74.3	0	10,350	95.3	95.3
(0 = no, 1 = yes)	1	1,914	25.7	100.0	1	516	4.8	100.0

Notes: ⋆ This sample was the first we collected. At that time, we exclusively considered five-door cars to control for convertibles, which could not be distinguished from other three-door vehicles. In the other samples an upgrade of the parsing software allowed us filter directly for convertibles.

Table 4. *Summary Statistics - Audi A4*[⋆]

	Panel A. Series B6 (Est. Period: 10/2001 - 10/2004)				Panel B. Series B7 (Est. Period: 04/2005 - 09/2007)			
Variable	Obs.	Mean	Min	Max	Obs.	Mean	Min	Max
price [EUR]	4,055	11,986	4,000	21,900	6,938	18,249	6,900	39,990
mileage [km]	4,055	111,448	9,729	233,550	6,938	90,143	1,000	234000
power [kW]	4,055	106	66	169	6,938	115	66	294
totalage [months]	4,055	68	51	87	6,938	34	16	45
Dummies	Value	Frequency	Percent	Cum.	Value	Frequency	Percent	Cum.
diesel	0	2,655	48.1	48.1	0	1,440	20.8	20.8
(0 = petrol, 1 = diesel)	1	2,864	51.9	100.0	1	5,498	79.2	100
auto gearbox	0	3,936	71.3	71.3	0	4,735	68.3	68.3
(0 = manu., 1 = auto)	1	1,583	28.7	100.0	1	2,203	31.8	100
cruise control	0	3,877	70.3	70.3	0	3,236	46.6	46.6
(0 = no, 1 = yes)	1	1,642	29.8	100.0	1	3,702	53.4	100
seat heating	0	2,641	47.9	47.9	0	2,126	30.6	30.6
(0 = no, 1 = yes)	1	2,878	52.2	100.0	1	4,812	69.4	100
all-wheel-drive	0	4,892	88.6	88.6	0	5,752	82.9	82.9
(0 = no, 1 = yes)	1	627	11.4	100.0	1	1,186	17.1	100
sun roof	0	4,719	85.5	85.5	0	5,974	86.1	86.1
(0 = no, 1 = yes)	1	800	14.5	100.0	1	964	13.9	100
leathertrim	0	4,545	82.4	82.4	0	5,337	76.9	76.9
(0 = no, 1 = yes)	1	974	17.7	100.0	1	1,601	23.1	100
metallic paint	0	938	17.0	17.0	0	1,710	24.7	24.7
(0 = no, 1 = yes)	1	4,581	83.0	100.0	1	5,228	75.4	100
private seller	0	4,276	77.5	77.5	0	6,506	93.8	93.8
(0 = no, 1 = yes)	1	1,243	22.5	100.0	1	432	6.2	100

Notes: ⋆ The Audi A4 production line includes no three-door version.

Table 5. *Summary Statistics - Opel Astra**

	Panel A. Series G (Est. Period: 01/2000 - 12/2003)				Panel B. Series H (Est. Period: 05/2004 - 10/2007)			
Variable	Obs.	Mean	Min	Max	Obs.	Mean	Min	Max
price [EUR]	2,825	5,332	1,750	9,950	8,358	10,820	3,900	19,900
mileage [km]	2,825	108,788	2,732	206,000	8,358	63,969	1,000	206,000
power [kW]	2,825	69	48	188	8,358	80	55	177
totalage [months]	2,825	85	61	108	8,358	35	15	56
Dummies	Value	Frequency	Percent	Cum.	Value	Frequency	Percent	Cum.
diesel	0	2,104	74.5	74.5	0	3,817	45.7	45.7
(0 = petrol, 1 = diesel)	1	721	25.5	100.0	1	4,541	54.3	100.0
five-door	0	463	16.4	16.4	0	760	9.1	9.1
(0 = three, 1 = five)	1	2,362	83.6	100.0	1	7,598	90.9	100.0
auto gearbox	0	2,574	91.1	91.1	0	7,878	94.3	94.3
(0 = manu., 1 = auto)	1	251	8.9	100.0	1	480	5.7	100.0
cruise control	0	2,649	93.8	93.8	0	1,007	12.1	12.1
(0 = no, 1 = yes)	1	176	6.2	100.0	1	7,351	88.0	100.0
seat heating	0	2,653	93.9	93.9	0	7,472	89.4	89.4
(0 = no, 1 = yes)	1	172	6.1	100.0	1	886	10.6	100.0
sun roof	0	2,679	94.8	94.8	0	8,202	98.1	98.1
(0 = no, 1 = yes)	1	146	5.2	100.0	1	156	1.9	100.0
leathertrim	0	2,738	96.9	96.9	0	7,968	95.3	95.3
(0 = no, 1 = yes)	1	87	3.1	100.0	1	390	4.7	100.0
metallic paint	0	520	18.4	18.4	0	1,390	16.6	16.6
(0 = no, 1 = yes)	1	2,305	81.6	100.0	1	6,968	83.4	100.0
private seller	0	2,395	84.8	84.8	0	7,832	93.7	93.7
(0 = no, 1 = yes)	1	430	15.2	100.0	1	526	6.3	100.0

Notes: * None of the observations for Opel Astra included an all-wheel drive.

Table 6. *The impact of year discontinuities on price*
Dependent variable: Car Price

	(1)	(2)	(3)	(4)
Indicator for...				
...older than 2007	-1671.8***	1779.0***	304.3***	-359.1***
	(74.4)	(145.4)	(85.0)	(68.5)
...older than 2006	-1685.6***	36.5	-274.2***	-543.1***
	(56.1)	(99.8)	(58.8)	(51.6)
...older than 2005	-2904.8***	-1454.9***	-552.1***	156.1*
	(50.9)	(84.7)	(62.9)	(62.3)
...older than 2004	-1336.5***	-63.0	-594.9***	38.4
	(52.5)	(103.1)	(83.8)	(81.9)
...older than 2003	-1084.6***	-94.1	-383.1***	-344.6***
	(52.3)	(94.6)	(81.8)	(78.2)
...older than 2002	-1190.9***	-286.0***	-116.2	53,2
	(46.5)	(80.6)	(84.0)	(89.0)
...older than 2001	-1345.4***	-570.1***	-85.9	-116.5
	(41.6)	(79.9)	(95.8)	(102.4)
...older than 2000	-1311.6***	-721.5***	107.2	-355.5***
	(37.2)	(88.0)	(106.2)	(99.9)
...older than 1999	-638.3***	-215.3**	-47.2	-566.7***
	(36.8)	(71.6)	(157.7)	(146.5)
5th-order age polynomial		X	X	X
Controls for car features			X	X
Car model fixed effects				X
R-squared	0,4816	0,4987	0,8355	0,8866
N	106076	106076	84491	73401

Notes: Robust standard errors in brackets. *** indicates $p < 0.001$; ** indicates $p < 0.01$; * indicates $p < 0.05$.

Table 7. *The impact of mileage discontinuities on price*
Dependent variable: Car Price

	(1)	(2)	(3)	(4)
Indicator for...				
...more than 10K km	-722.9***	-383.4***	-431.4***	-545.0***
	(103.2)	(107.4)	(74.7)	(77.6)
...more than 20K km	-280.5**	102.7	-207.7**	-379.8***
	(88.7)	(88.7)	(68.7)	(71.2)
...more than 30K km	-878.8***	-491.9***	-1184.5***	-1211.8***
	(94.1)	(94.1)	(73.0)	(71.6)
...more than 40K km	-307.1**	69.2	-461.0***	-399.2***
	(96.3)	(96.3)	(72.0)	(70.0)
...more than 50K km	-369.6***	-15.7	-115.3	-39.4
	(90.2)	(90.2)	(64.6)	(64.5)
...more than 60K km	167.7	504.3***	-73.3	-47.2
	(86.7)	(86.7)	(61.7)	(61.6)
...more than 70K km	-318.4***	17.0	-22.3	-17.3
	(79.9)	(79.9)	(59.6)	(59.8)
...more than 80K km	-118.5	192.5*	-41.1	-50.6
	(74.7)	(74.7)	(56.9)	(56.8)
...more than 90K km	-440.7***	-135.0	-116.7*	-34.0
	(70.8)	(70.8)	(55.0)	(55.6)
3rd-order mileage polynomial		X	X	X
Controls for car features			X	X
Car model fixed effects				X
R-squared	0,4929	0,4936	0,8264	0,8775
N	106076	106076	84491	73401

Notes: Robust standard errors in brackets. *** indicates $p < 0.001$; ** indicates $p < 0.01$; * indicates $p < 0.05$.