GOVERNANCE AND THE EFFICIENCY OF ECONOMIC SYSTEMS

GESY

Discussion Paper No. 505

Price discontinuities in an online market for used cars

Florian Englmaier*
Arno Schmöller**
Till Stowasser***

* University of Munich
** TWS Partners
*** University of Munich

March 2015

Financial support from the Deutsche Forschungsgemeinschaft through SFB/TR 15 is gratefully acknowledged.

Price discontinuities in an online market for used cars*

Florian Englmaier† Arno Schmöller‡ Till Stowasser§

March 15, 2015

Abstract

We use more than 63,000 datapoints from a German used car market website to document systematic and substantial price drops at vintage (= year of first registration) thresholds and 10,000 km odometer marks. The latter finding replicates the findings in Lacetera *et al.* (2012), whereas the first dimension cannot be analyzed with their US data because only German cars have such legally mandated and regulated "birthdates". Hence we have the unique opportunity to study the presence of coarse information processing within the same dataset and decision problem but across two separate domains. We document that discontinuities in these two domains are of comparable size. While Lacetera *et al.* (2012) explain their result with a left-digit bias in the processing of numerical information, vintage discontinuities cannot be explained by this. We propose a slightly more general model of information prominence and availability bias to accommodate our findings.

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 participants at the 2013 IIOC conference in Boston, the VfS 2013 Meeting in Düsseldorf, the 2014 German IO comittee meeting in Hamburg and in particular Maximilian Breu, Matthias Dischinger, Thomas Fackler, Nicola Lacetera, Devin Pope, Stephen Ryan, Klaus Schmidt, Peter Schwardmann, Justin Sydnor, and Steve Tadelis for their helpful comments and suggestions. Ines Helm provided excellent research assistance. This research was partially funded through DFG grant *SFB/TR-15*. A previous version of the paper was circulated as: Florian Englmaier and Arno Schmöller "The Evaluation of Complex Goods: Evidence from Online Car Sales".

†University of Munich, florian.englmaier@econ.lmu.de

‡TWS Partners, info@schmoeller.org

§University of Munich, till.stowasser@econ.lmu.de

1 Introduction

Economic theory suggests that a rational agent should incorporate all relevant information when making a decision. However, at least since Simon (1955), economists have proposed models that relax this strong assumption. In these models, individuals simplify complex decisions, for example, by processing only a subset of information. Moreover, recent empirical research convincingly documents that consumers fail to efficiently process the available relevant information and instead rely on heuristic evaluation rules.[1] In particular, Lacetera *et al.* (2012) use (literally) millions of datapoints from US used car auctions, find systematic and substantial price drops at 10,000 mile odometer marks, and explain this pattern with a model of inattention based on left digit bias.[2]

We use comprehensive field data on used car offers from the German website *mobile.de*, one of Europe's largest online vehicle marketplaces and are able to replicate the Lacetera *et al.* (2012) findings. But the German context allows us to pursue an additional line of inquiry, which helps us to gain a better understanding of the mechanism driving our results: In Germany, cars have a legally mandated official documentation record that makes the date of first-registration verifiable information. Importantly, the "model year" concept is not used for German cars. Instead of issuing a new model each year, German manufacturers produce a given model generation without significant changes for a period of several years. Hence, we have the unique possibility to study a second dimension where coarse information processing might play a role within one data set and a single decision problem.

We document strong threshold effects on prices at year changes in the date of first-registration. All else equal, the price differential between 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. We are documenting our results by implementing a regression discontinuity design and results are robust to applying differing sets of controls, controlling for polynomials of lower or higher order than suggested by the Akaike Information Criterion test, or using log-linearized data. A linear approximation of a limited attention model, as suggested by DellaVigna (2009) or Lacetera *et al.* (2012), suggests that the inattention parameter, capturing the unexplained

[1] See, e.g. Lee and Malmendier (2011) and Brown *et al.* (2010) on internet auctions or Chetty *et al.* (2009) and Finkelstein (2009) on taxes and tolls.

[2] For example, cars with odometer values between 59,000 and 59,999 miles are sold only slightly cheaper than cars with odometer readings between 58,000 and 58,999 miles but the price drop to the 60,000-60,999 bin is substantially larger.

price drops at thresholds is between 0.3 and 0.4 across both dimensions, i.e. comparable in size and well within the range of parameter estimates documented in the prior literature (see below).

Lacetera *et al.* (2012) suggest that discontinuities regarding odometer readings can be reconciled with a model of *left-digit bias* in information processing. However, in a literal interpretation, this cannot explain the discontinuities between a car fist registered in December (e.g., 12/200**4**) and January (e.g., 01/200**5**). We argue that a slightly more general model of information prominence and availability bias can accommodate our findings. Such a model would be readily applicable to a wide set of domains where underlying continuous characteristics are classified in discrete categories, e.g., classifications of French wine or ratings of financial assets, where "inefficient" pricing discontinuities might also exist.

The paramount role of information provision in online markets is underlined by Lewis (2011). Tadelis and Zettelmeyer (2015) document it for the used car market also studied by Lacetera *et al.* (2012). Limited attention has also been documented for purchase decisions in other markets. For instance, Lee and Malmendier (2011) 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 process. Pope (2009) shows that patients strongly react to changes in response to changes in (coarse) rankings of hospitals while they ignore more informative measures of hospital quality. In a similar vein, the degree of salience of taxes appears to affect consumption behavior. For example, Chetty *et al.* (2009) conduct a field experiment at a grocery store and find that posting tax-inclusive prices reduces demand. Finkelstein (2009) shows that reduced salience of road tolls (caused by the introduction of electronic toll collection systems) leads to higher tolls. Analyzing stock market data, Gilbert *et al.* (2012) 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 inattention 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 data. Section 3 presents the graphical and regression analysis for vintage and mileage discontinuities. Section 4 presents a simple model to rationalize our results. Section 5 presents linear approximations of structural parameters to capture otherwise unexplained price drops. Section 6 concludes and the Appendix collects all Figures and Tables.

2 Data

2.1 Data source: The mobile.de website

For the purpose of this study, we collected detailed information on more than 63,000 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 intermediary 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*".[3] 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 allows to filter for makes, models, and a number of other basic details. An advanced search form provides a large additional set of filter options.

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. "01/2000", meaning that the car was first registered in January of 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 selected cars.

`-- Include Figure 2 about here. --`

A typical profile page for an offered car, which is accessed by clicking the respective search result, is depicted in Figure 3.

[3]Source: http://cms.mobile.de/en/company/portrait_mobile.html; last accessed: May 1, 2013

-- Include Figure 3 about here. --

Our data does not come from an auction market and here, instead of final prices, we observe seller's asking prices only. These may be subject to negotiation before a car is sold. Note, however, that for this to be a concern, we would need differential deviations from the asking price for pre- and post-threshold cars, which seems implausible. Moreover, there are important reasons to believe that the asking price is a good proxy for the final price in this market. First, *mobile.de* offers the seller the 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 car model the market for used cars is highly competitive. Moreover, cars within each model generation are very close substitutes. Assuming that the stated asking price reflects the seller's willingness to accept an offer, 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. Similarly, 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) or List (2004) the endowment effect is unlikely to apply. In fact, the vast majority of offers in our sample comes from commercial rather than private sellers. 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 (2009) document that the sellers' reserve prices in a similar, but distinct, online-auctions market are – as suggested by theory – continuous functions of valuations and hence similarly determined as the sales prices; i.e. from an evaluation of the individual attributes. Our intuition is that the same applies here. For simplicity, we use the term "price" to refer to the stated asking prices throughout the paper.

2.2 Sample composition

Our data includes information for the most widespread car models from the four leading German car manufacturers[4], all ranked among the top ten of Germany's vehicle population according to the Federal Office for Motor Vehicles (Kraftfahrtbundesamt/KBA).[5] Specifically, we collected information on 14,780 Volkswagen (VW) Golf (KBA-rank 1), 10,841 Opel Astra (KBA rank 2), 18,470 BMW 3 series (KBA rank 4), 14,219 Audi A4 (KBA rank 7), and 5,030 Mercedes Class A (KBA rank 9), all advertised as accident-free and with their first registration-dates between 01/2000 and 12/2007.[6]

[4]Due to limited resources, we focused on the best-selling model of each of the four biggest German car makers.

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

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

During the summer of 2009 , we collected data on cars whose date of first registration falls between December 2007 and January 2000. This means that the youngest car in our sample is 20 months old, whereas the oldest car has an age of 116 months (roughly 9 and a half years). This selection excludes a certain type of used car from our sample, referred to as "Jahreswagen", i.e. cars given for roughly a year to employees as part of their compensation package.[7]

`-- Include Table 1 about here. --`

Because the introduction of a new model generation affects prices substantially, we can only retrieve meaningful estimates for the influence of car attributes if we accurately control for model revisions. This requires detailed knowledge of the exact dates of the respective market launches. For the four models considered in our sample this information is readily available through either the manufacturers' websites, the so-called Schwacke-List (http://schwacke.de), or the Deutsche Automobil Treuhand (http://www.dat.de).[8] For an overview of model updates that occurred in our sample, see Table 1 and note that there is no clustering of updates around year changes.

However, controlling for these model updates is not trivial because we only know when factories switched production from old to new models but cannot observe if a given car in our sample is truly a new model or rather an old model that has been sitting at the dealer's lot for a few months. Furthermore, some makes (e.g., BMW) do not introduce all their model variants at the same date but in a sequential way (e.g., the station wagon is introduced 8 months after the sedan). For our main specifications we classify a car as having undergone a model update if it was first registered more than 3 months after the factories switched production. In Sections 3.2 and 3.4 we document that our results are robust to alternative definitions of these indicators.

2.3 Summary statistics

Table 2 provides an overview of the main variables and the corresponding summary statistics. In general, the value of an individual car from a specific model series depends on numerous factors. 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, seat-heating, or cruise control. Besides the asking prices

[7] The market of Jahreswagen is special and these cars are rather different (e.g. they are usually pricy vehicles with high-end additional features) from the average car offered on a market for second-hand cars. Results – which are available upon request – that include these Jahreswagen show that our results remain stable while discontinuities are unsystematic in this market segment.

[8] These 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; this renders the discontinuities we document in our data even more remarkable.

and the month and year of first registration, we therefore collected a large number of car features to control for quality differences. To measure their impact on the price of the car, we create dummy variables for the respective car features.

`-- Include Table 2 about here. --`

The information on the month and year of the first registration is stored in the variables $fr\_month \in [1,12]$ and $fr\_year \in [2000,2007]$, respectively. For our empirical analysis, we combine them to the variable $totalage \in [1,96]$, which captures the precise age of a car measured in months:

$$totalage \equiv 12 \cdot (2007 - fr\_year) + (13 - fr\_month),$$

where the car's age is normalized relative to the youngest car in the sample (i.e. a registration date 12/2007).

When inspecting Table 2 note that almost 90% of offers comes from professional car dealers, as indicated by the dummy *private_seller* being equal to 0.13. As one would expect, *price* is strongly negatively correlated 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 other extras are positively related to the price of a car. We also find that *totalage* and *mileage* co-move at a degree of $\rho = 0.77$. Even though, in general, collinearity among the explanatory variables can be problematic, our sample size is sufficiently large to produce precise parameter estimates.

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 the color black. We find that the prices are indeed somewhat responsive to different colors. To facilitate presentation, however, we will omit the coefficients for the color-dummies in the regression tables that follow.

`-- Include Figure 4 about here. --`

`-- Include Figure 5 about here. --`

`-- Include Figure 6 about here. --`

Next, consider the age distribution of the cars, which is depicted in Figure 4. We find some fluctuation across registration months but our sample contains a sufficient number of observations for each first registration-date in the estimation period. The distribution of Mileage increases up to 30,000 km and is pretty much evenly distributed beyond 30,000 km (Figure 5). Finally, the distribution of prices is somewhat right skewed, but approximately normally distributed (Figure 6).

3 Empirical analysis

3.1 Vintage discontinuities

Graphical analysis We begin the empirical analysis by plotting the raw price data as a function of car age. In Figure 7, each dot shows the average asking price for all cars first registered in a given month of a given year starting December 2007 and counting backwards until January 2000.

`-- Include Figure 7 about here. --`

As one would expect, average prices decrease with increasing age. Within each year, monthly average prices decline almost linearly, but there are discontinuities between years (Figure 7). These patterns are systematic and substantial for all cars in our sample. The picture becomes even clearer once we control for heterogeneity in the car population by plotting adjusted residuals (average prices after controlling for a 5th-order age polynomial, mileage, horsepower, model update, and other car features) in Figure 8.

`-- Include Figure 8 about here. --`

Regression analysis The graphical analysis suggests the existence of systematic price discontinuities at year changes for the month of first registration. To augment these results, we perform regression analysis to obtain numerical estimates for the observed price discontinuities. Our identification strategy is based on a regression discontinuity (RD) design; see Lee and Lemieux (2010) for an overview.

To identify the discontinuities at vintage thresholds, we estimate the regression equation

$$p_i = \alpha_v + \sum_{y=1}^{7} \beta_v D_{yi}^{v} + f(a_i) + \mathbf{X_i'}\gamma_v + \epsilon_{vi}, \tag{1}$$

where p_i is the car's asking price, α_v is a constant and $\mathbf{X_i'}$ is a vector of observable car characteristics. $f(\cdot)$ is a polynomial function of car age, a_i and is supposed to capture the continuous relationship between price and car age.[9] We also include seven dummy variables, D_{yi}^{v} that indicate whether a car has crossed a given year threshold. The corresponding β_v coefficients are the parameters of interest, as they measure the discontinuous difference in prices – conditional on the exact car age – when comparing cars that were registered in different years. The intuition for this approach is that the coarse registration year information should have no additional impact on car values, once the much finer information in a_i is accounted for.

[9] In our main specification we use a fifth-order polynomial to control for age. This specific functional form was chosen based on the Akaike Information Criterion test. Our results are robust to using other polynomial orders, as discussed in Section 3.2.

-- Include Table 3 about here. --

Table 3 presents regression results for the specification described above. Column (1) merely controls for a fifth-order age polynomial and the full set of threshold indicators and provides estimates of the price discontinuities before accounting for heterogeneity in the car population. Given the important role of (in this specification uncontrolled) model updates, these somewhat unsystematic results should be interpreted with caution. In Columns (2) through (5) we increase the number of control variables. Column 2 adds controls for basic car features, such a mileage, horse power, etc. which substantially increases R^2 and affects significance, size, and in some cases even the sign of the coefficients. In Column (3) we augment the control vector by information on model updates while Column (4) includes model fixed effects. Column (5) is our preferred specification as it simultaneously controls for all of the aforementioned characteristics. Once heterogeneity in car features is accounted for, almost all of the coefficients of interest are significantly negative and for 2006, 2005, 2002, and 2001 rather large (on average > EUR 400 relative to an average price of EUR 13,000). These results closely mirror the graphical analysis above. Even though the pattern is not perfect in terms of significance, we conclude that there is strong evidence for sizable and systematic negative price discontinuities upon passing a year threshold even after controlling for the exact age and a host of observable characteristics.

3.2 Robustness – Vintage discontinuities

3.2.1 Potential empirical pitfalls

Do other car features change at vintage thresholds? To assure that our results are not driven by differing car populations around the thresholds, we check whether January and December cars are comparable with respect to their average mileage (see Figure 9), their average horsepower (see Figure 10), their fuel type (see Figure 11), and the composition of seller types (dealer/private) (see Figure 12). These raw data plots show somewhat erratic and certainly unsystematic patterns, which suggests that the discontinuities are not driven by changes in the market composition at year's end.

-- Include Figure 9 about here. --

-- Include Figure 10 about here. --

-- Include Figure 11 about here. --

-- Include Figure 12 about here. --

Are there discontinuities in the density of cars at thresholds? Another typical robustness check in RD design settings consists of testing for manipulative sorting around the thresholds. i.e., whether the price drops could be explained by supply "shocks". Note that ex-ante we should not expect to find systematic patterns here: Because the registration date of the car is an inherent feature that does not change (and cannot be changed) over time, there is no incentive (or even an opportunity) to sell the car prior to crossing a threshold (in contrast to the case of mileage, below). Indeed, while the distribution looks jumpy and is subject to seasonality, there are no discontinuities at the year thresholds (see Figure 13). These data follow the seasonal pattern of first registrations that the Federal Office for Motor Vehicles (Kraftfahrtbundesamt) records since 1970.

`-- Include Figure 13 about here. --`

Are there discontinuities at Placebo thresholds? In another important consistency check we perform placebo tests by creating indicators for artificial year thresholds and testing whether these are associated with discontinuities as well. Even though in a few cases we do get significant coefficients for the Placebo thresholds, they are not systematic, neither in size nor in sign. Results for three different placebo tests that move the vintage threshold to the other three end-of quarter months – March, June, and September – can be found in Table 4.

`-- Include Table 4 about here. --`

3.2.2 Robustness to changes of the main specification

Log-linearization Prices have a long right tail, hence log-linearization seems appropriate. Results are robust to log-linearizing prices; see Table 5. In fact, results from this specification look even cleaner and more systematic than those from our main specification; which seems particularly reassuring given the above mentioned long right tail of prices.

`-- Include Table 5 about here. --`

Controlling for model updates For our main specification we classify a car as having undergone a model update if its date of first registration was more than 3 months after factories switched production. Our results are robust to varying definitions of model update dummies. This suggests that the potential measurement error when assigning the model generation is unlikely to be harmful. Table 6 collects the results for the following dummy definitions that have been used for the robustness analysis[10]:

- D1: Model generation dummy (i.e., fixed-effects for each model generation) imposing no insecurity: We treat all cars registered 1 month or more after the official model switch as being a new model.
- D2: Model generation dummy (i.e., fixed-effects for each model generation) with 5-months insecurity windows: Like D1 but we treat model status of cars registered within the three months after a model switch as "unknown". Effectively, these cars are not used to identify our model.
- D3: Model update dummy, imposing no insecurity: Takes on the value of 1 if a model update happened in a given month.
- D4: Model update dummy, with 5-months insecurity windows: Like D3 but the three months after the introduction are also labeled as a model update month.

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

Higher- and lower-order age polynomials A key identifying assumption of any RD design is that the continuous relationship between the forcing variable (here: car age) and the outcome variable (here: car price) be adequately captured by the polynomial function $f(\cdot)$. Since this assumption is inherently untestable, the literature – see e.g., Lee and Lemieux (2010) – stresses that results should at least be robust to varying polynomial functions in order to be credible. As can be seen in Table 7 this is indeed the case, as our key results stay stable even when lower and higher ordered polynomials than those endorsed by information criteria are used instead.

`-- Include Table 7 about here. --`

3.3 Mileage discontinuities

We now turn to the other car feature that could potentially be associated with price discontinuities: the odometer reading as stated in the sales offer. Note that this information is self-reported and hence less reliable than the officially regulated and mandated registration dates. In fact, we observe apparent rounding in the mileage data and there is some indication that sellers strategically declare lower odometer readings, especially when the true value has just crossed a psychologically relevant 10.000 km threshold. We discuss this below. Despite these potential measurement error concerns, we are able to replicate the findings from Lacetera *et al.* (2012).

[10]We have also experimented with smaller and larger insecurity windows (3 to 6 months) but omitted these results to save space. Results are very similar.

3.3.1 Graphical analysis

Again, we begin by plotting the raw price data as a function of car mileage. In Figure 14, each dot shows the average asking price for cars in a 1,000 km mileage bin. There is a dot for the average price of cars with 1,000 through 1,999 km, then a dot for cars with 2,000 to 2,999 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 km band, average prices decline quite smoothly but there are systematic (albeit small) drops at the 10,000 km marks. A similar and more systematic picture emerges in Figure 15 when plotting adjusted residuals (average prices after controlling for a 2nd-order mileage polynomial, car age, horsepower, model updates, and other car features) over odometer readings; note that the trendlines, while helpful to visualize the downward price trend across and within bins, blur the price drops. To see these, focus on the triangles indicating average (adjusted) residual car prices measured in 1,000km bins around the 10,000km thresholds.

`-- Include Figure 14 about here. --`

`-- Include Figure 15 about here. --`

With no apparent explanation for the importance of 10,000 km thresholds, this result suggests, following Lacetera *et al.* (2012), a role for heuristic decision making along the mileage dimension, as well.

3.3.2 Regression analysis

To complement this visual evidence, we again turn to regression analysis to establish numerical estimates for these price discontinuities. As before, we implement RD designs where the dependent variable is the car price as stated on the website. The regression equation takes the following form:

$$p_i = \alpha_m + \sum_{k=1}^{10} \beta_m D_{ki}^m + g(o_i) + \mathbf{X_i'}\gamma_m + \epsilon_{mi}, \tag{2}$$

where $g(\cdot)$ is a polynomial function of odometer readings, o_i capturing the continuous relationship between price and car mileage.[11] To measure the impact of crossing a given 10,000 km threshold, we include ten dummy variables, D_{ki}^m. The corresponding β_m coefficients are the parameters of interest, as they measure the discontinuous difference in prices – conditional on the exact (stated) odometer reading – when a car crosses a 10,000 km mark.

[11]In our main specification we use a second-order polynomial to control for mileage. This specific functional form was chosen based on the Akaike Information Criterion test. The robustness with respect to other polynomials is discussed in Section 3.4.

-- Include Table 8 about here. --

Even though the graphical evidence looks less clear-cut than was the case for the vintage discontinuities, the corresponding regression results (see Table 8) look reassuring. As before, Column (1) controls only for the mileage polynomial and the full set of indicator variables. As we move from Column (2) to Column (5), control for car heterogeneity becomes more encompassing. When accounting for all observable characteristics in Column (5), all threshold coefficients are significantly negative and sizeable, suggesting economically meaningful price discontinuities at 10,000 km marks.

3.4 Robustness – Mileage Discontinuities

3.4.1 Potential empirical pitfalls

Do other car features change at 10,000 km thresholds? To assure that our results are not driven by differing car populations around the thresholds, we check whether cars around the 10,000 km thresholds are comparable with respect to their average mileage (see Figure 16), their average horsepower (see Figure 17), their fuel type (see Figure 18), and the composition of seller types (dealer/private; see Figure 19). As before with the vintage thresholds, the raw data plots show erratic and unsystematic patterns. This again suggests that the discontinuities are not driven by systematic changes in the market composition at thresholds.

-- Include Figure 16 about here. --

-- Include Figure 17 about here. --

-- Include Figure 18 about here. --

-- Include Figure 19 about here. --

Are there discontinuities in the density of cars at thresholds? In Figure 20 we find indirect evidence that sellers do at least believe that the mileage thresholds are important for car prices or, alternatively, the likelihood of them being sold: Clearly, cars are brought to the market just before the odometer passes a 10,000 km threshold. Note that the observed pattern can be reconciled with sellers misreporting and declaring slightly lower odometer readings so that cars barely fall into the more attractive mileage bin. This alternative explanation, however, would have very similar implications. Importantly, a simple excess supply story (whether virtual or real) cannot account for the observed discontinuities because this should depress rather than inflate prices for cars in this crowded market segment.

-- Include Figure 20 about here. --

Are there discontinuities at Placebo thresholds? For our placebo test we convert odometer readings from kilometers to US miles and check for discontinuities in the converted data at 10,000 mile thresholds. Given that Germany exclusively uses the metric system, US miles thresholds should be irrelevant even if heuristics play a role in decision making. As is apparent from Table 9, and as to be expected, all but the very first 10,000 mile placebo threshold are associated with insignificant coefficients.

`-- Include Table 9 about here. --`

3.4.2 Robustness to changes of the main specification

Log-linearization Mileage discontinuities are not as robust as those for the price drops in the vintage domain. Other than in the analysis above, log-linearization is not innocuous, but even reverses results for some 10,000 km thresholds (i.e. the estimates here are significantly positive); see Table 10. This inversion of some signs for log-linearized data is puzzling. Discontinuities in the mileage dimension have been convincingly documented to be systematic and sizeable by Lacetera *et al.* (2012). As compared to their data, our data probably suffer from more measurement error. Hence we are certainly not inclined to call their findings into question but rather attribute the lack of robustness in this dimension to our more limted and somewhat less precise – w.r.t. mileage information – data.

`-- Include Table 10 about here. --`

Controlling for model updates Again, our results are robust to varying definitions of model update dummies. Hence also for mileage, measurement error when assigning the model generation is unlikely to be harmful. Table 11 collects the results for various dummy definitions; see Section 3.2.2 for the definition of alternative model update dummies.

`-- Include Table 11 about here. --`

Higher- and lower-order age polynomials As with log-linearization, results are somewhat sensitive to using polynomials of orders higher than 3 (while orders of 1, 2, and 3 are fine); see Table 12. However, specifications that accommodate polynomials of orders higher than 3 appear to suffer from multicollinearity and STATA does not provide the F-Test-statistic. I.e., as long as the econometric model is not misspecified, results are in line with our predictions.

`-- Include Table 12 about here. --`

3.5 A Horserace

Finally, we implement a horserace specification by controlling for mileage and vintage thresholds at the same time. Results in Table 13 are, once more, encouraging and suggest that both discontinuities exist independent of one another. Not only do effects in either dimension survive, especially the vintage discontinuities appear slightly more pronounced.

`-- Include Table 13 about here. --`

4 A simple model

In light of the rather sophisticated behavior when jointly evaluating more than a dozen car features, the main finding we document in our data seems even more intriguing. If people are careful enough to compare numerous details of a car's attribute vector, why do they systematically pay too little attention to the valuable information captured in the month of first registration or the exact mileage? Though they do not completely disregard the impact of precise age as indicated by the continuous decline within a vintage (or mileage bin), they apparently fail to fully recognize the connection to subsequent or previous vintages (or mileage bins). Our intuition is that individuals evaluate cars relative to the average car from an easily accessible "comparison sample", i.e., the same vintage or mileage bin, while the more relevant comparison group consists of cars of similar age or mileage, irrespective of the vintage or mileage bin they belong to.

As an illustration, consider a prospective buyer who wants to evaluate a car with given attributes first registered in 12/2006. Other things equal, she "should" compare prices of cars of comparable age, such as a six-months window (all cars registered between 09/2006 and 03/2007). However, this is in addition to potential pre-existing biases, also hampered by the website's search mask as neither *mobile.de*'s simple nor its advanced search allow to adjust inquiries for the precise month of first registration. Hence, in order to obtain the desired information, the agent would have to screen a substantially higher number of offers, namely the entire universe of cars registered in 2006 and 2007. This suggests that one possible source of the discontinuities lies within the design of *mobile.de*'s search interface that differentially makes information for calendar year based reference groups of cars more directly available and prominent. A comparable argument applies to the 10,000 km brackets.

To capture this in a formal model, consider a risk-neutral agent j who wants to evaluate a particular car $i = (y_i, m_i, X_i)$, where $y_i \in \{2000, \dots, 2008\}$ denotes its vintage, $m_i \in \{0, \dots, 12\}$ the month of first registration and X_i all other attributes of the car.[12]

[12] We set out the model in terms of vintage discontinuities, but it can be applied equivalently to mileage discontinuities.

Normalize by $a_i = 12 \cdot (2008 - y_i) + (13 - m_i)$ the total age in months. For given values of $X_i = \overline{X}$, individual j's value estimate for car i dependent on its age attribute is described by the function $E_j[v_i] : (y_i, m_i) \to \mathbb{R}^+$.[13] More specifically, let

$$E_j[v_i] := (1-\theta) \cdot \overline{v}_y^j + \theta \cdot v_{a_i}^j,$$

where $\overline{v}_y^j$ is the value of an average car in age-group y, and $v_{a_i}^j$ denotes her precise value of car i. For simplicity, assume that $\overline{v}_y^j$ 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 θ captures the weight of the exact (actually correct) valuation in the eventual decision utility. 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 θ on her true value for car i and thereby obtains a more precise estimate.[14] Formally, assume that the convex weight θ captures the following properties $\theta = 0$ corresponds to fully coarse thinking, $\theta \in (0,1)$ corresponds to partially coarse thinking, and $\theta = 1$ corresponds to fully rational thinking.

First, consider the case where information is sufficiently hard to access such that $\theta = 0$. Then buyer j's valuation will reflect the average value $\overline{v}_y^j$. Second, suppose that $\theta < 1$. If $v_{a_i}^j > \overline{v}_y^j$, the buyer 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 would be willing to 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 affect the posted final price. Third, if information is easily available, agent j will fully learn her precise value, i.e. $\theta = 1$. These cases are illustrated in Figure 21.

`-- Include Figure 21 about here. --`

For higher θ, the estimates of any individual agent j should become more accurate in the sense that they become closer to her precise value $v_{a_i}^j$. Our data are thus consistent with a model that includes partially coarse thinking, $0 < \theta < 1$, which suffice to cause discontinuities between two consecutive vintages. Obviously, a similar reasoning can be applied to explain the price drops around 10,000 km odometer marks since the user interface only offers to look for cars within coarse pre-defined mileage bins.

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

[14] One way to understand the underlying rationale is to assume that for any θ the agent solves an underlying optimal search problem, which implicitly determines the extent to which she learns $v_{a_i}^j$ and the weight the exact valuation has in her decision problem.

5 Estimating the extent of the friction

The key parameter in the above model is $(1-\theta)$ as it captures the friction that occurs when consumers update their price evaluation with increasing car age or odometer readings. The friction can be interpreted as caused by search costs or – as in DellaVigna (2009) or in Lacetera *et al.* (2012) – as an inattention parameter. In either case, the higher $(1-\theta)$, the larger are the discontinuities when crossing a threshold. In this section we generate linear approximations for this friction parameter.

To do so we assume a true linear value function, v_i; after inspecting Figure 7 and Figure 14 we deem this to be not an entirely unreasonable approximation, in particular if one abstracts from the first 10,000 km bracket. The size of the estimated price discontinuity at a vintage (or, alternatively, 10,000 km) threshold is approximately equal to $\alpha * (1-\theta) * \Delta$ where α is the slope of the value function, and Δ is the width of the bracket, i.e., one year or 10,000 km. Geometrically, $(1-\theta)$ measures the fraction of the discontinuous (unexpected) price reduction that occurs at the thresholds. Imposing linearity allows us to approximate the friction parameter, as it forces the discontinuities to be constant across thresholds. While this specific functional form is unlikely to be the best fit to our data, it is a simplification commonly used in the literature; see, e.g., DellaVigna (2009) or Lacetera *et al.* (2012).

Fitting this model with our data, we obtain estimates of 0.32 and 0.39 for friction in the vintage and in the mileage domain, respectively. These results suggest that approximately 30-40% of the depreciation that a car experiences occurs discontinuously at year changes and 10,000 km thresholds. These numbers are comparable in size, suggesting that the effect of coarse thinking is comparable across domains at least in our data and both estimates fall well within the range of what has been documented for inattention parameters in the literature; see DellaVigna (2009).

6 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 coarse information processing. Despite the large monetary stakes involved, our findings suggest that people in this market systematically fail to aggregate the 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 first-registration-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 justifiable, 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. Hence our setting gives us the unique possibility to study the role of coarse thinking across two domains within the same decision problem. Further quantitative analysis documents that the extent of discontinuities across these two domains is comparable in size.

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 finding relating to odometer thresholds, 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 distorted choice behavior due to differential prominence or availability of information that is capable of explaining both 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.

This research may be extended in several ways. Our suggested model of prominence or availability of information is applicable to any domain where underlying continuous characteristics are classified in discrete categories, from French wine to ratings of financial assets. Future studies can evaluate whether pricing discontinuities also exist in these contexts.

In his seminal contribution to information economics, Akerlof (1970) uses 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 (or even verifiable) characteristics. Consumers 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. 17

Brown, J., Hossain, T. and Morgan, J. (2010). Shrouded attributes and information suppression: Evidence from the field. *Quarterly Journal of Economics*, **125(2)**, 859–876. 1

Chetty, R., Looney, A. and Kroft, K. (2009). Salience and taxation: Theory and evidence. *American Economic Review*, **99(4)**, 1145–1177. 1, 2

DellaVigna, S. (2009). Psychology and economics: Evidence from the field. *Journal of Economic Literature*, **47(2)**, 315–372. 1, 16

Englmaier, F. and Schmöller, A. (2009). Determinants and effects of reserve prices in hattrick auctions. *Working Paper*. 4

Finkelstein, A. (2009). Ez tax: Tax salience and tax rates. *Quarterly Journal of Economics*, **124(3)**, 969–1010. 1, 2

Gilbert, T., Kogan, S., Lochstoer, L. A. and Ozyildirim, A. (2012). Investor inattention and the market impact of summary statistics. *Management Science*, **58** (2), 336–350. 2

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. 4

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. 4

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, 10, 11, 13, 16, 17

Lee, D. S. and Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature*, **48(2)**, 281–355. 7, 10

Lee, Y. H. and Malmendier, U. (2011). The bidder's curse. *American Economic Review*, **101(2)**, 749–787. 1, 2

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

List, J. A. (2004). Neoclassical theory versus prospect theory: Evidence from the marketplace. *Econometrica*, **72** (2), 615–625. 4

Pope, D. G. (2009). Reacting to rankings: Evidence from "america's best hospitals". *Journal of Health Economics*, **Vol. 28(6)**, pp. 1154–1165. 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. 4

Simon, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics*, **Vol. 69(1)**, pp. 99–118. 1

Tadelis, S. and Zettelmeyer, F. (2015). Information disclosure as a matching mechanism: Theory and evidence from a field experiment. *American Economic Review*, **105** (2), 886–905. 2

A Figures and Tables

Figure 1. *Website www.mobile.de*

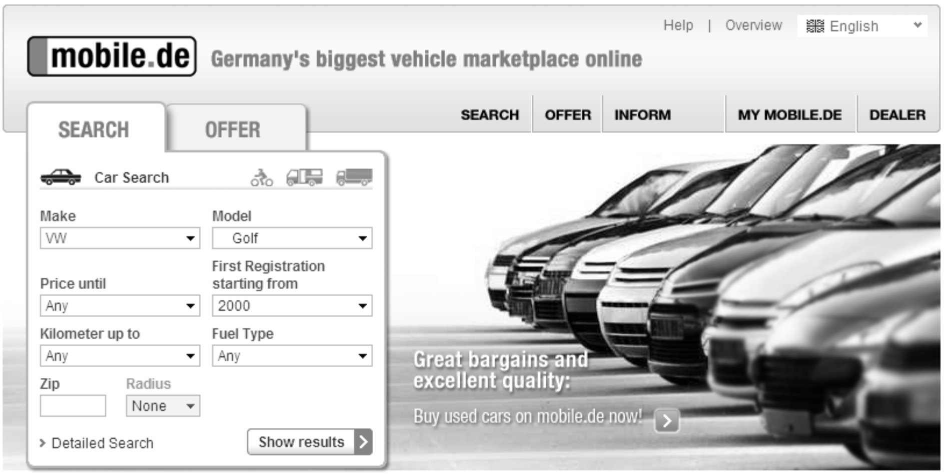


Figure 2. *Search Results Listing – www.mobile.de*

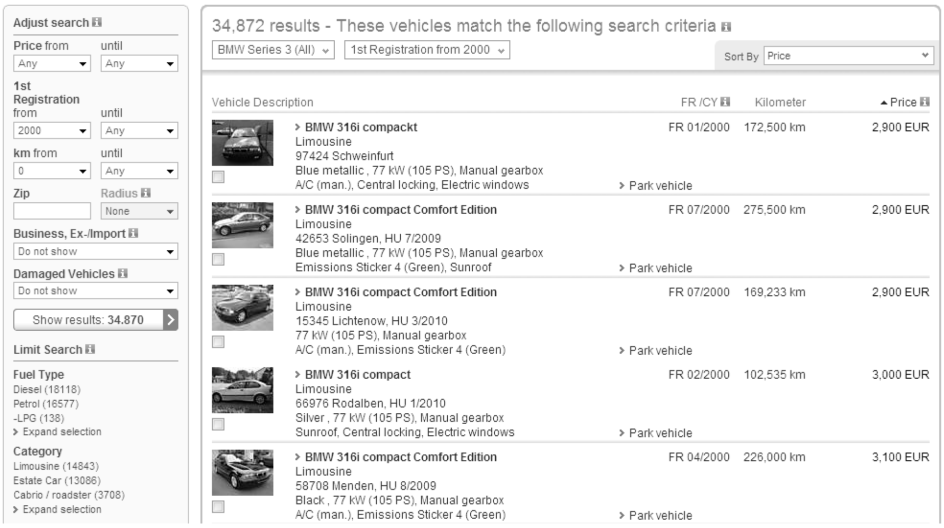


Table 1. *Overview of Model Generations*

Make & Model	Name of Series		Production period
Audi A4	B6	(limousine)	10/2000 - 11/2004
		(estate)	09/2001 - 11/2004
	B7	(limousine)	11/2004 - 11/2007
		(estate)	11/2004 - 03/2008
BMW 3	E46		04/1998 - 11/2004
	E90	(limousine)	12/2004 - 09/2008
		(estate)	06/2005 - 09/2008
Mercedes A Class	168		09/1997 - 09/2004
	169		10/2004 - 04/2012
Opel Astra	G		02/1998 - 01/2004
	H		02/2004 - 10/2007
VW Golf	IV		10/1997 - 09/2003
	V		10/2003 - 07/2008

Notes: The table shows the production periods of the sub-series of all respective makes and models in our sample.

Table 2. *Summary Statistics*

Variable	N	Mean	StDev	Min	Max
Price (in EUR)	63,340	13,209	5,909	1,750	48,890
Mileage (in km)	63,340	87,338	51,629	1,000	1,499,000
Total age (in months*)	63,340	39.14	25.82	1	96
Horsepower (in PS)	63,289	97.26	28.94	44	309
Indicators:					
Diesel engine	63,322	0.58	0.49	0	1
Automatic transmission	63,018	0.21	0.40	0	1
Metallic paint	63,340	0.79	0.41	0	1
Air condition	63,340	0.97	0.17	0	1
Leather trim	63,340	0.16	0.37	0	1
Airbag	63,340	0.49	0.50	0	1
Power windows	63,340	0.96	0.21	0	1
Sunroof	63,340	0.18	0.38	0	1
Four-wheel drive	63,340	0.05	0.22	0	1
Seat heating	63,340	0.45	0.50	0	1
Cruise control	63,340	0.43	0.50	0	1
Private seller	63,340	0.13	0.34	0	1

Notes: *Total age in months measured relative to December 2007.

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

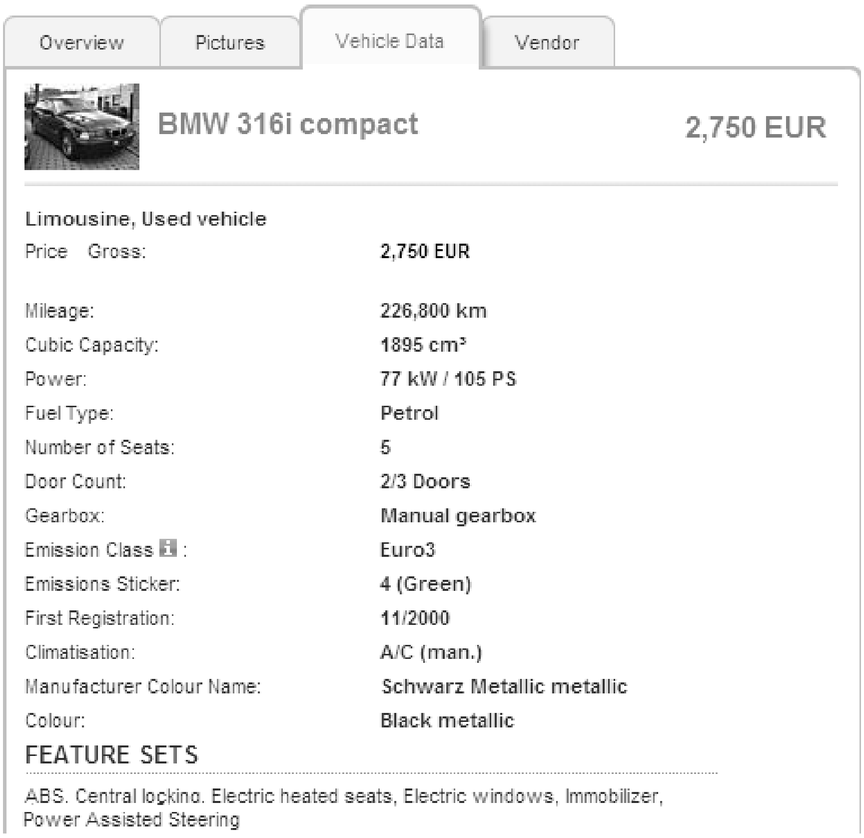
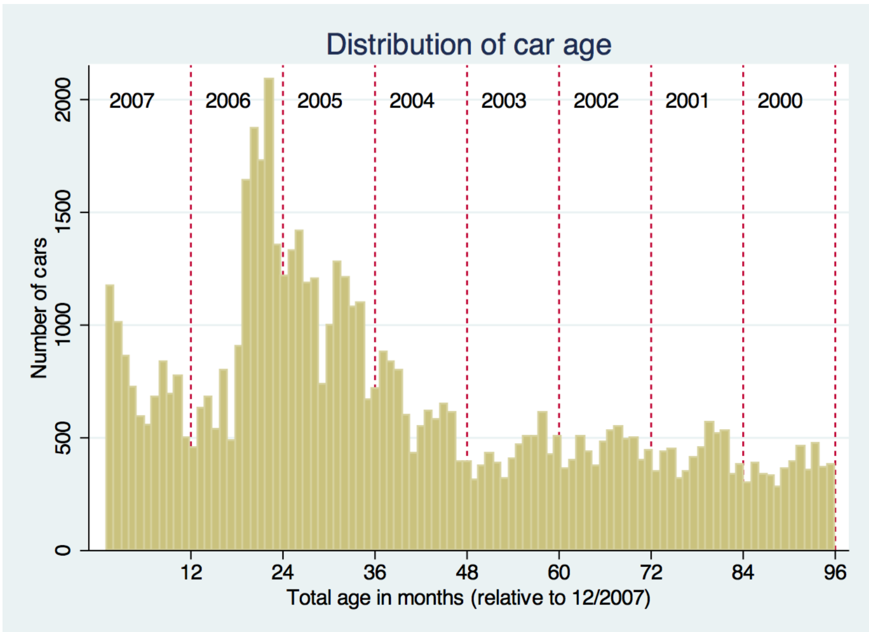
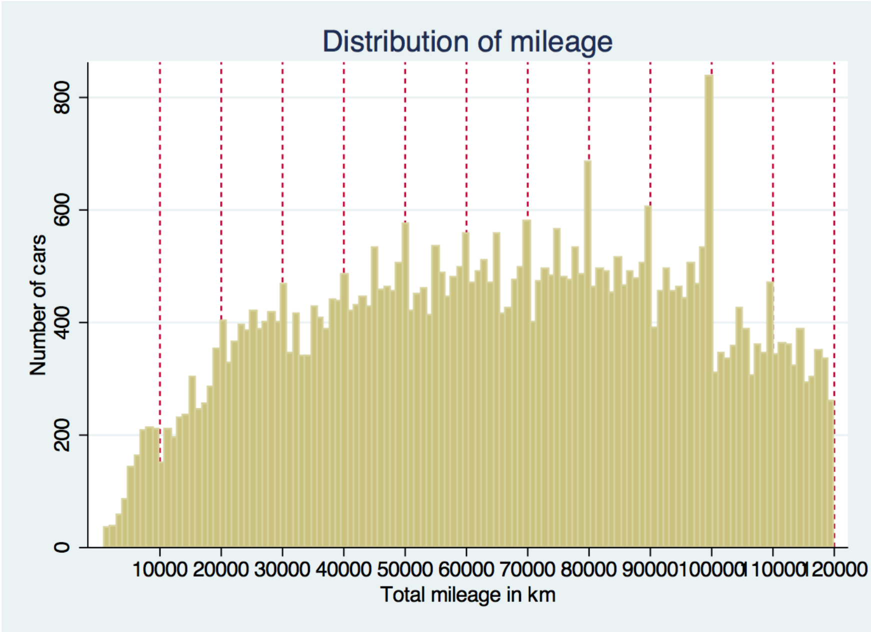


Figure 4. *Distribution of Car Age*



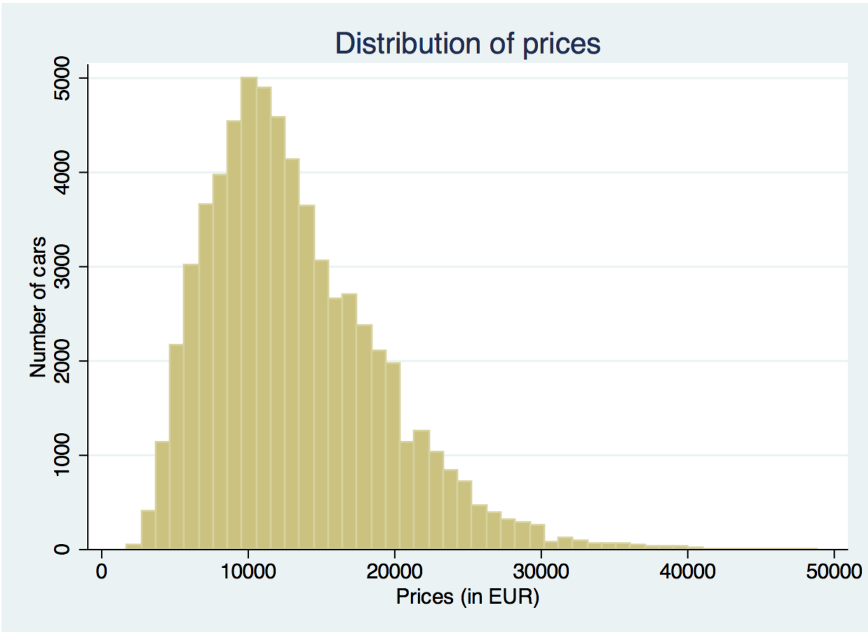
Notes: The figure plots the distribution of car age, measured relative to date of first registration 12/2007, on a monthly basis. Vertical lines in the graph indicate each year.

Figure 5. *Distribution of Car Mileage*



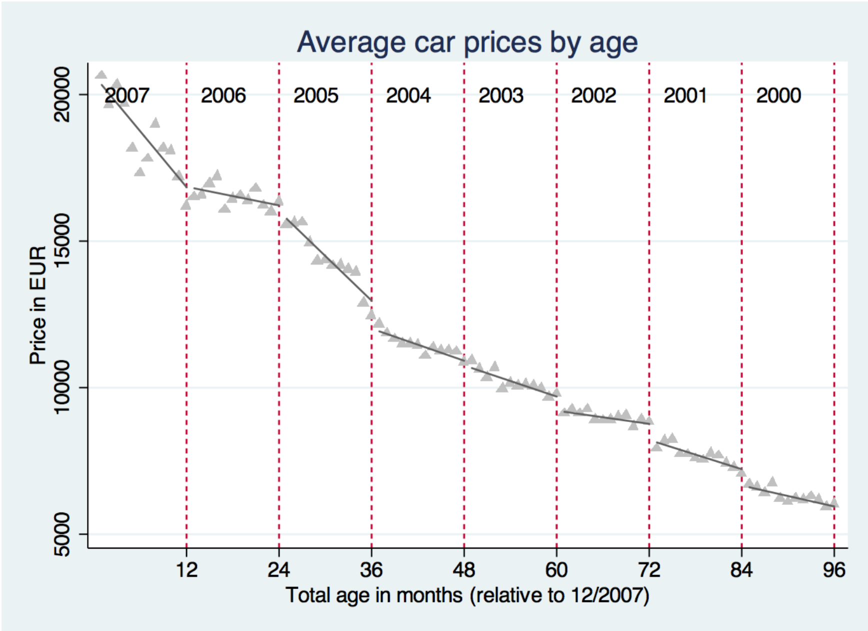
Notes: The figure plots the distribution of car mileage, measured in 1,000km bins. Vertical lines in the graph indicate each 10,000km threshold.

Figure 6. *Distribution of Car Prices*



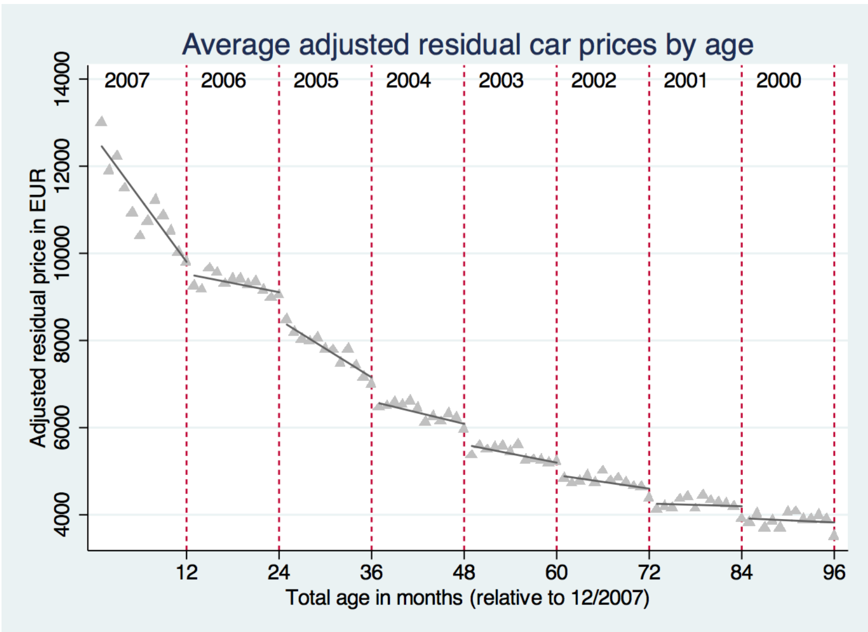
Notes: The figure plots the distribution of car ask prices, measured in EUR 1,000 increments.

Figure 7. *Avg. Car Prices by Age (monthly averages)*



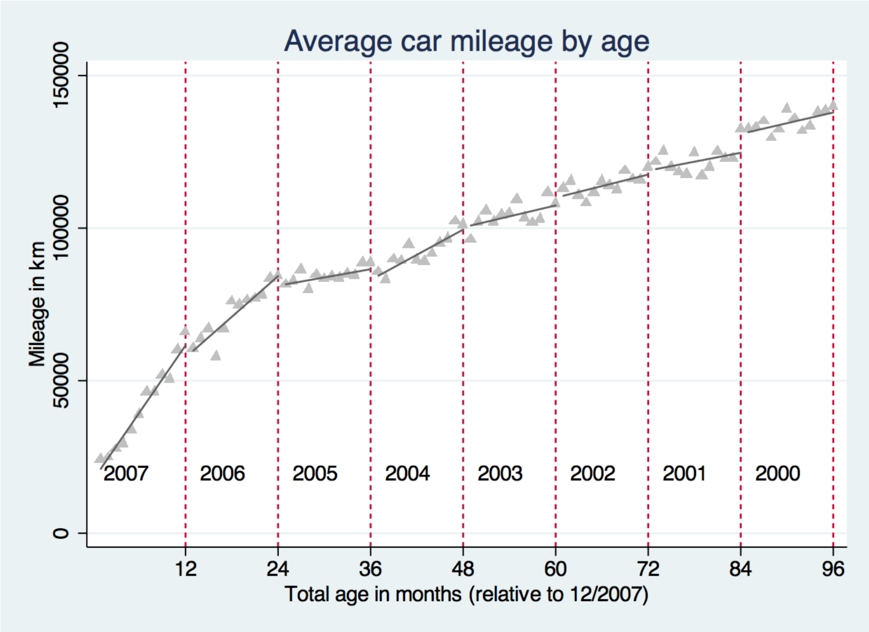
Notes: The figure plots the average (raw) car prices, measured on a monthly basis. Age is measured relative to the date of first registration 12/2007. Vertical lines in the graph indicate each year thresholds.

Figure 8. *Avg. Adj. Residual Car Prices - all cars*



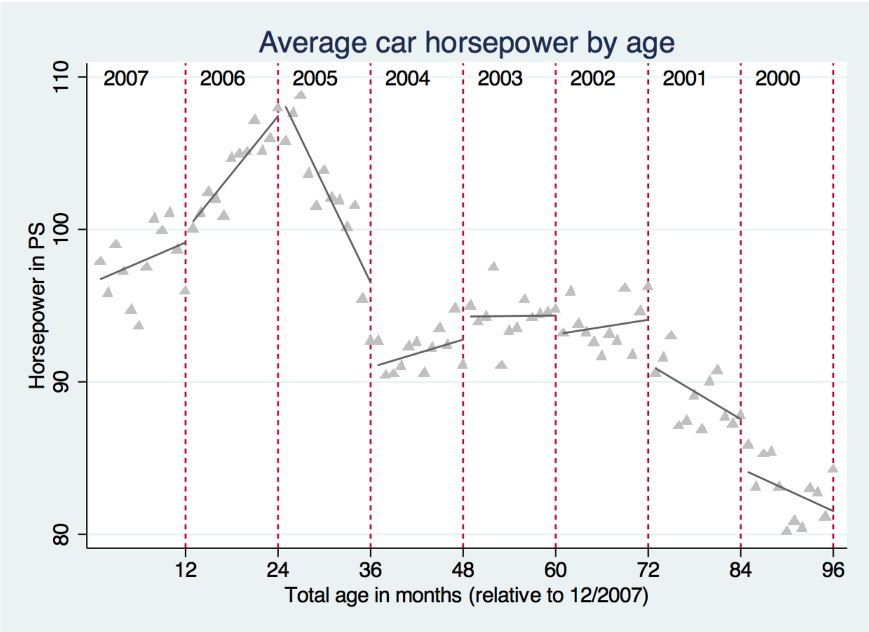
Notes: The figure plots the average (adjusted) residual car prices after controlling for car characteristics and measured on a monthly basis. Age is measured relative to the date of first registration 12/2007. Vertical lines in the graph indicate each year thresholds.

Figure 9. *Avg. Car Mileage by age (monthly averages)*



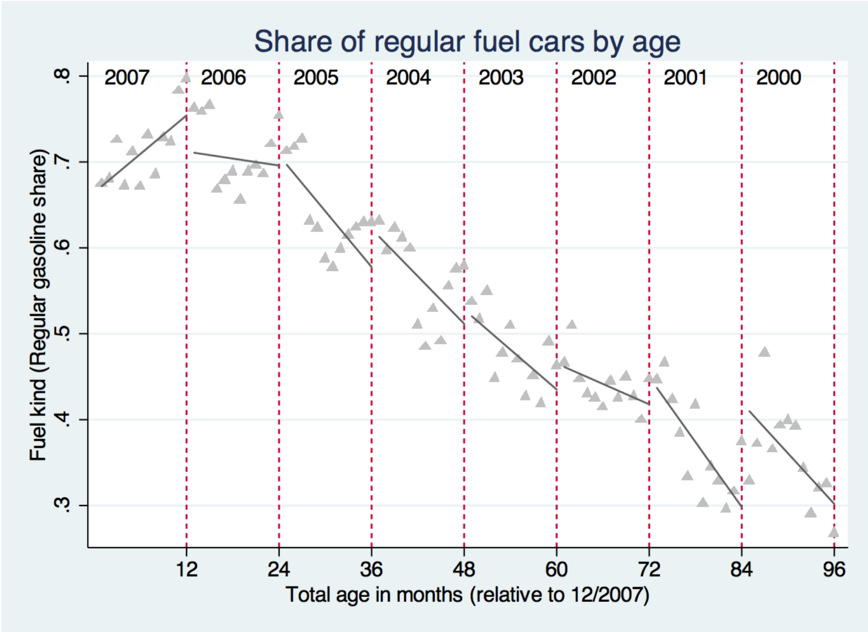
Notes: The figure plots the average mileage measured on a monthly basis. Age is measured relative to the date of first registration 12/2007. Vertical lines in the graph indicate each year thresholds.

Figure 10. *Avg. horsepowers by age (monthly averages)*



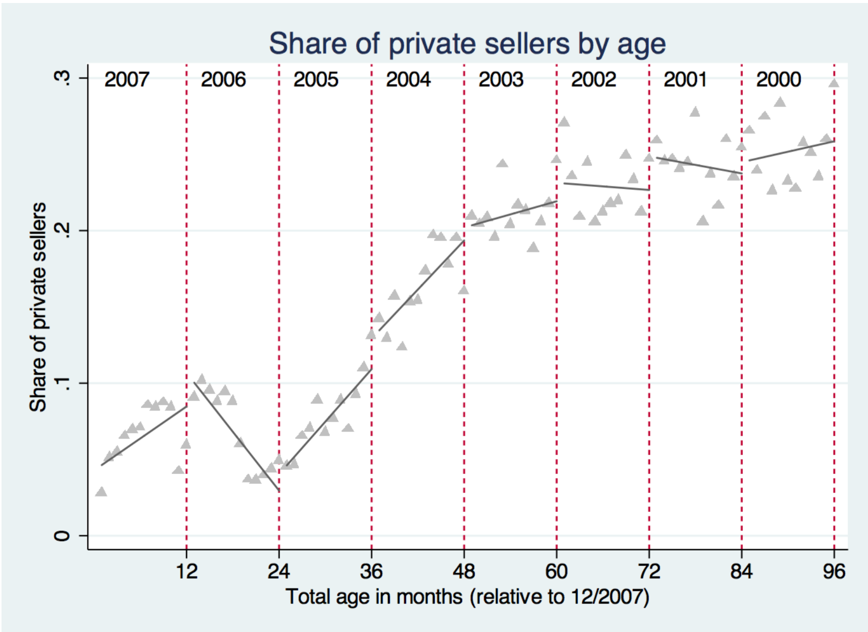
Notes: The figure plots the average horsepowers of traded cars measured on a monthly basis. Age is measured relative to the date of first registration 12/2007. Vertical lines in the graph indicate each year thresholds.

Figure 11. *Share of regular gasoline cars on offer by age (monthly averages)*



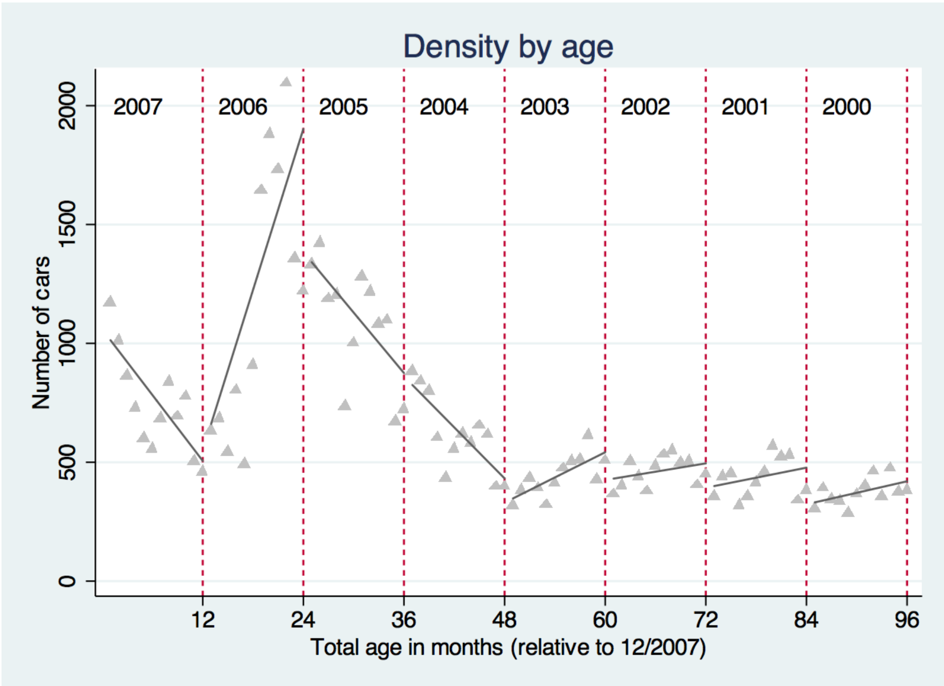
Notes: The figure plots the average share of regular gasoline (= non–Diesel) cars on sale measured on a monthly basis. Age is measured relative to the date of first registration 12/2007. Vertical lines in the graph indicate each year thresholds.

Figure 12. *Avg. Share of Private Sellers by age (monthly averages)*



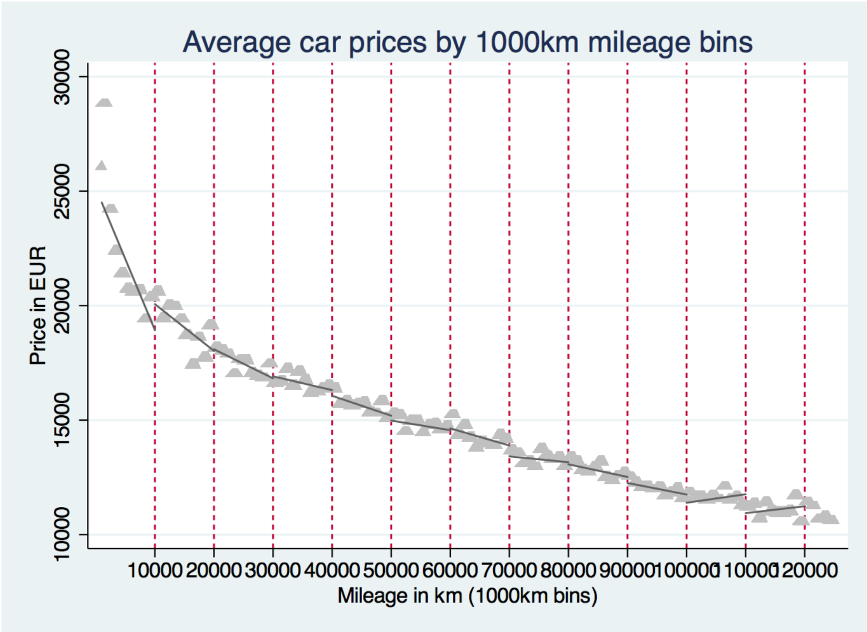
Notes: The figure plots the average share of cars sold by private sellers measured on a monthly basis. Age is measured relative to the date of first registration 12/2007. Vertical lines in the graph indicate each year thresholds.

Figure 13. *Avg. Number of Cars offered by age (monthly averages)*



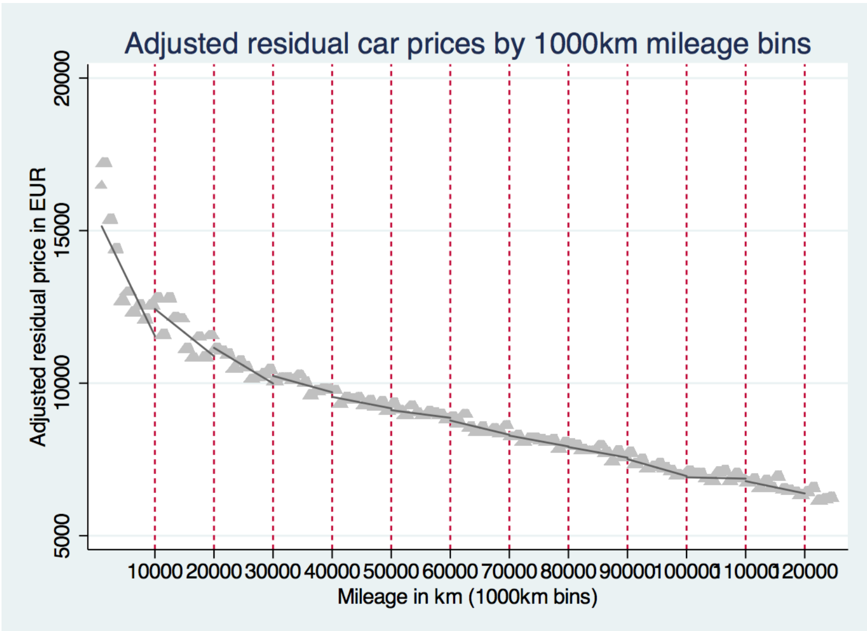
Notes: The figure plots the average number of cars on offer measured on a monthly basis. Age is measured relative to the date of first registration 12/2007. Vertical lines in the graph indicate each year thresholds.

Figure 14. *Distribution of Car Prices by Mileage (1,000Km bins)*



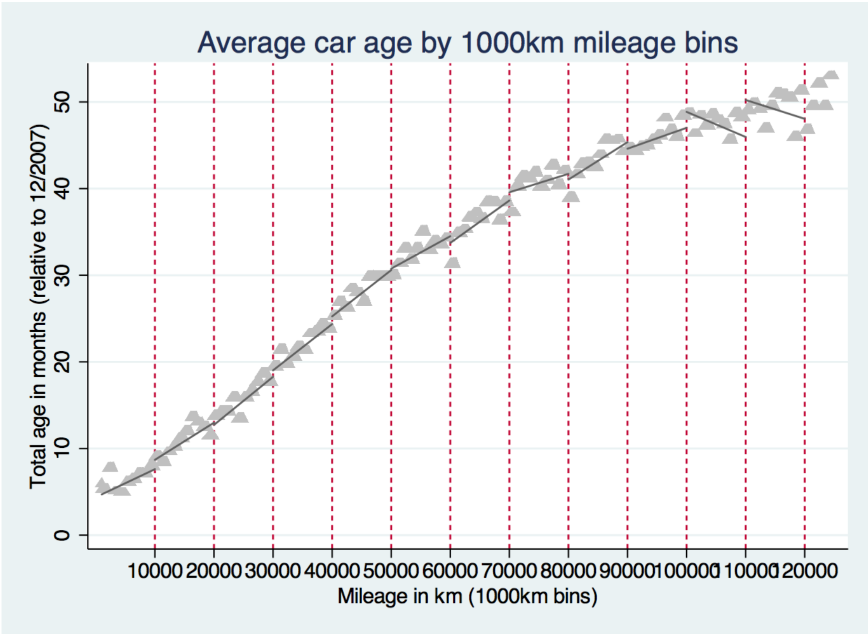
Notes: The figure plots the average (raw) car prices measured in 1,000km bins. Vertical lines in the graph indicate each 10,000km threshold.

Figure 15. *Avg. Adj. Residual Car Prices by Mileage (1,000Km bins)*



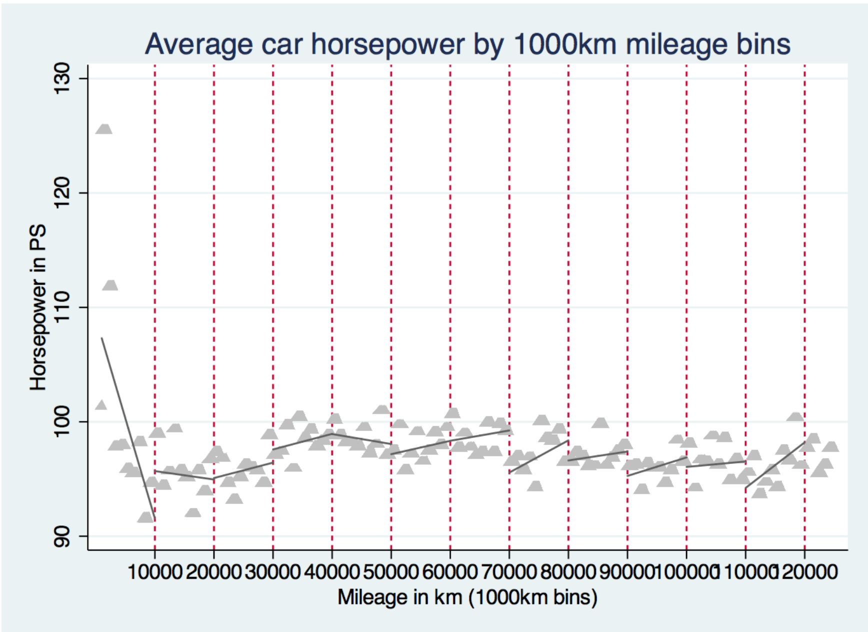
Notes: The figure plots the average (adjusted) residual car prices after controlling for car characteristics and measured in 1,000km bins. Vertical lines in the graph indicate each 10,000km threshold.

Figure 16. *Avg. Car Age by Mileage (1,000km bin averages)*



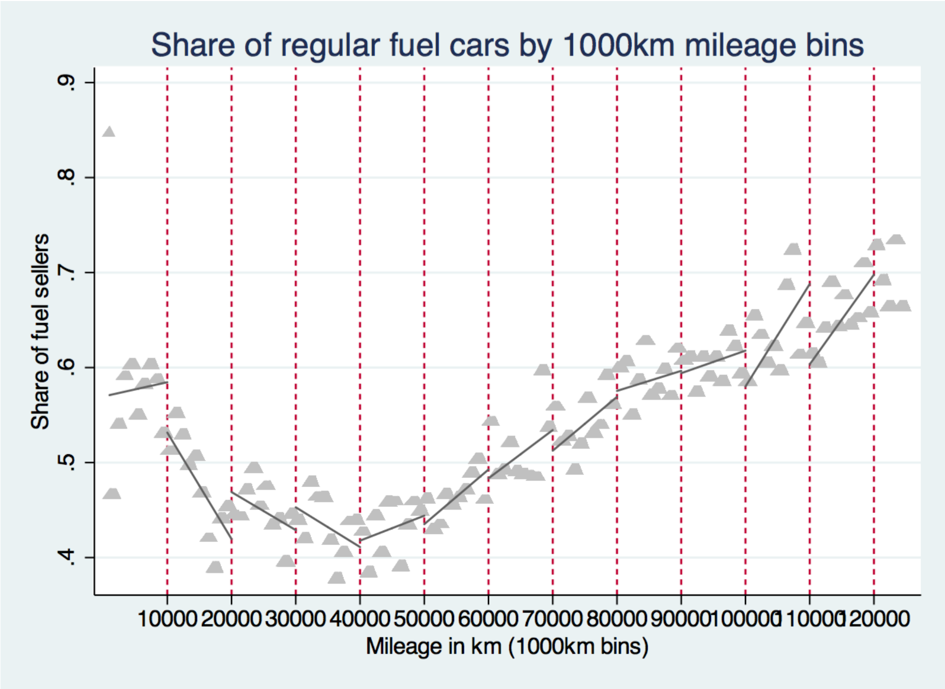
Notes: The figure plots the average age of offered cars measured in 1,000km bins. Vertical lines in the graph indicate each 10,000km threshold.

Figure 17. *Avg. horsepowers by Mileage (1,000km bin averages)*



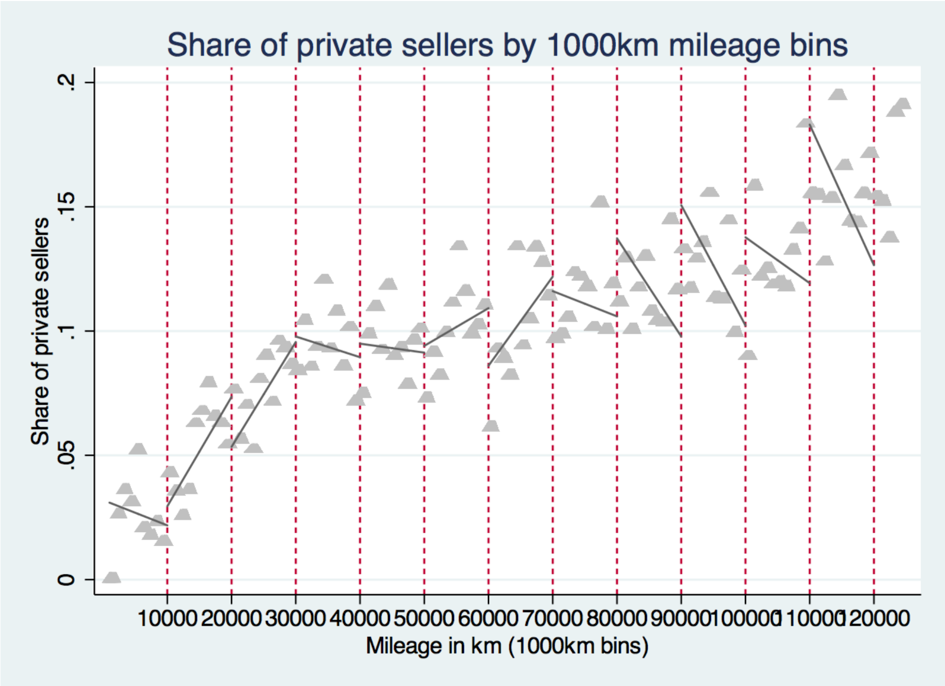
Notes: The figure plots the average horsepowers of traded cars measured in 1,000km bins. Vertical lines in the graph indicate each 10,000km threshold.

Figure 18. *Share of regular gasoline cars on offer by Mileage (1,000km bin averages)*



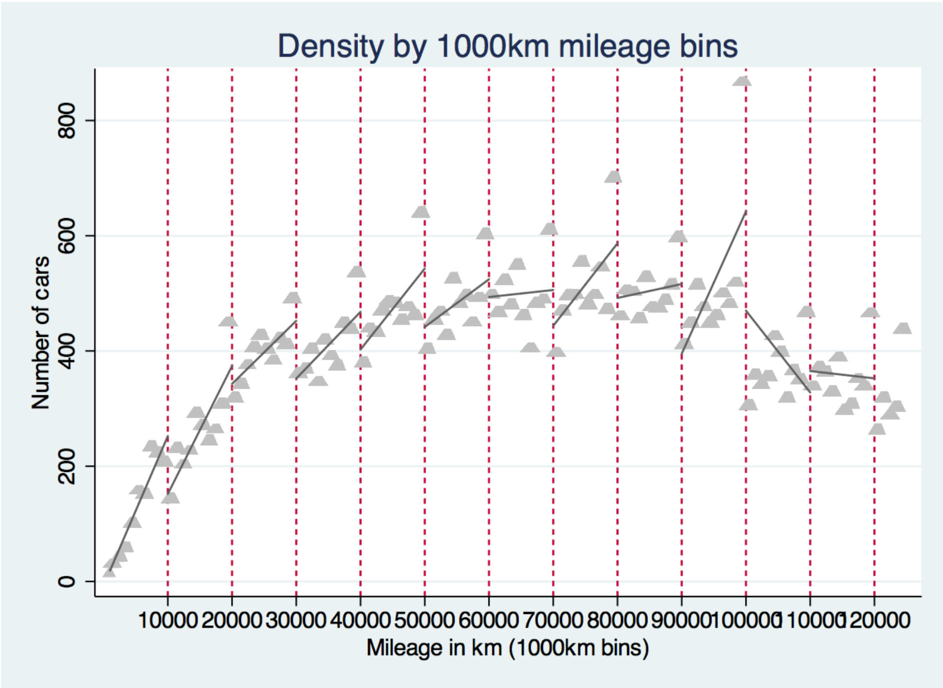
Notes: The figure plots the average share of regular gasoline (= non–Diesel) cars offered measured in 1,000km bins. Vertical lines in the graph indicate each 10,000km threshold.

Figure 19. *Avg. Share of Private Sellers by Mileage (1,000km bin averages)*



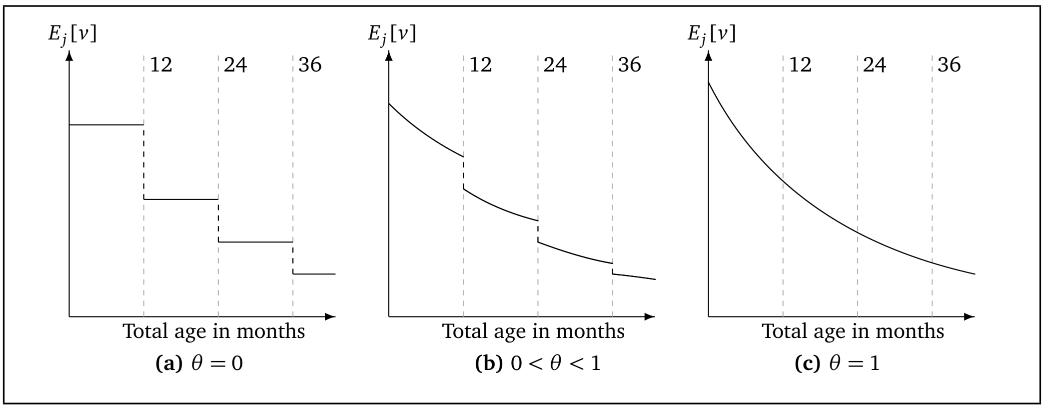
Notes: The figure plots the average share of offered cars sold by private sellers measured in 1,000km bins. Vertical lines in the graph indicate each 10,000km threshold.

Figure 20. *Avg. Number of Cars offered by Mileage (1,000km bin averages)*



Notes: The figure plots the average number of of offered cars measured in 1,000km bins. Vertical lines in the graph indicate each 10,000km threshold.

Figure 21. *Expected Valuation in Dependence of Search Costs*



Notes: The figure plots the theoretically predicted price distribution for cars. The (inattention or search) friction decreases from left to right panels. Vertical lines in the graph indicate year thresholds.

Table 3. *The impact of vintage discontinuities on price*

Dep. Variable: Car price	(1)	(2)	(3)	(4)	(5)
Indicator for...					
...older than 2007	430.7*	54.7	-203.0	-14.4	-221.4*
	(176.9)	(108.4)	(108.0)	(92.2)	(91.8)
...older than 2006	-501.8***	-613.7***	-858.4***	-580.5***	-780.3***
	(116.4)	(66.0)	(67.5)	(55.4)	(56.6)
...older than 2005	-1228.1***	-581.3***	-698.6***	-516.0***	-609.2***
	(96.0)	(63.5)	(63.6)	(54.6)	(54.9)
...older than 2004	576.9***	-51.0	-23.7	-30.7	-3.7
	(119.5)	(87.4)	(87.1)	(79.2)	(77.7)
...older than 2003	-42.0	-61.6	-129.6	105.6	44.3
	(100.5)	(77.2)	(76.8)	(69.5)	(68.9)
...older than 2002	-895.0***	-394.3***	-568.4***	-293.2***	-437.9***
	(107.2)	(88.7)	(88.5)	(82.7)	(83.5)
...older than 2001	-959.6***	-414.2***	-449.8***	-316.7***	-345.1***
	(94.1)	(93.7)	(93.5)	(90.9)	(91.7)
5th-order age polynomial	X	X	X	X	X
Controls for car features		X	X	X	X
Controls for model updates			X		X
Car model fixed effects				X	X
R-squared	0.4543	0.8148	0.8161	0.8576	0.8584
N	63,340	50,872	50,872	50,872	50,872

Notes: Identification strategy is based on a regression discontinuity (RD) design. The "...older than xxxx" dummy variables indicate whether a car has crossed a given year threshold. The corresponding coefficients are the parameters of interest, as they measure the discontinuous difference in prices – conditional on the exact car age – when comparing cars that were registered in different years. Robust standard errors in brackets. ***p<0.001; **p<0.01; *p<0.05

Table 4. *Robustness Analysis Vintage: Placebo month thresholds*

Dep. Variable: Car price	(1)	(2)	(3)
Indicator for...			
...older than 2007	-151.3	636.2***	-588.5**
	(110.4)	(146.0)	(181.0)
...older than 2006	79.7	455.0***	466.8***
	(55.5)	(75.9)	(90.3)
...older than 2005	-263.0***	-305.1***	-224.5***
	(57.3)	(59.6)	(57.4)
...older than 2004	229.6**	18.0	195.0**
	(77.4)	(73.9)	(65.3)
...older than 2003	225.5**	177.7	374.6***
	(73.6)	(94.1)	(93.7)
...older than 2002	-131.6	-8.5	185.2*
	(81.7)	(89.4)	(82.4)
...older than 2001	-313.6***	-207.2*	104.6
	(85.3)	(98.4)	(93.6)
5th-order age polynomial	X	X	X
Controls for car features	X	X	X
Controls for model updates	X	X	X
Car model fixed effects	X	X	X
Placebo month	March	June	September
R-squared	0.8586	0.8588	0.8588
N	50,872	50,872	50,872

Notes: Identification strategy is based on a regression discontinuity (RD) design. The "...older than xxxx" dummy variables indicate whether a car has crossed a given placebo year threshold. The corresponding coefficients are the parameters of interest, as they measure the discontinuous difference in prices – conditional on the exact car age – when comparing cars that were registered in different "years". All regressions also include the original non-placebo December months, for which estimates do not significantly change. Robust standard errors in brackets. ***p<0.001; **p<0.01; *p<0.05

Table 5. *Robustness Analysis Vintage: Log-linearization*

Dep. Variable: ln(Car price)	(1)	(2)	(3)	(4)	(5)
Indicator for...					
...older than 2007	0.024*	-0.008	-0.041***	-0.032***	-0.043***
	(0.010)	(0.006)	(0.006)	(0.005)	(0.005)
...older than 2006	-0.027***	-0.034***	-0.049***	-0.033***	-0.043***
	(0.008)	(0.004)	(0.004)	(0.003)	(0.004)
...older than 2005	-0.087***	-0.044***	-0.045***	-0.034***	-0.039***
	(0.008)	(0.005)	(0.005)	(0.004)	(0.004)
...older than 2004	0.017	-0.025**	-0.020**	-0.034***	-0.018**
	(0.011)	(0.007)	(0.007)	(0.006)	(0.006)
...older than 2003	-0.014	-0.017*	-0.020**	-0.008	-0.007
	(0.011)	(0.007)	(0.008)	(0.006)	(0.006)
...older than 2002	-0.087***	-0.052***	-0.060***	-0.046***	-0.053***
	(0.013)	(0.009)	(0.009)	(0.008)	(0.008)
...older than 2001	-0.111***	-0.075***	-0.076***	-0.070***	-0.071***
	(0.013)	(0.009)	(0.010)	(0.008)	(0.008)
5th-order age polynomial	X	X	X	X	X
Controls for car features		X	X	X	X
Controls for model updates			X		X
Car model fixed effects				X	X
R-squared	0.5340	0.8450	0.8416	0.8941	0.8945
N	63,340	50,872	50,872	50,872	50,872

Notes: Identification strategy is based on a regression discontinuity (RD) design. The "...older than xxxx" dummy variables indicate whether a car has crossed a given year threshold. The corresponding coefficients are the parameters of interest, as they measure the discontinuous difference in prices – conditional on the exact car age – when comparing cars that were registered in different years. Robust standard errors in brackets. ***p<0.001; **p<0.01; *p<0.05

Table 6. *Robustness Analysis Vintage: Varying controls for model updates*

Dep. Variable: ln(Car price)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Indicator for...								
...older than 2007	44.3	-56.5	126.4	-72.0	25.9	-33.8	-152.8	-158.6
	(108.3)	(89.4)	(108.4)	(90.7)	(107.7)	(91.8)	(108.2)	(91.9)
...older than 2006	-603.2***	-636.3***	-524.5***	-549.3***	-581.0***	-557.4***	-833.6***	-736.7***
	(65.7)	(54.0)	(66.0)	(54.9)	(65.8)	(55.3)	(66.9)	(56.0)
...older than 2005	-725.7***	-83.4	-505.9***	-352.0***	-751.0***	-643.2***	-547.9***	-492.1***
	(62.7)	(52.8)	(65.3)	(57.3)	(64.3)	(56.1)	(62.8)	(54.0)
...older than 2004	-122.1	70.8	51.1	53.3	-192.6*	-134.3	146.4	111.6
	(86.0)	(76.6)	(92.7)	(84.4)	(88.1)	(78.5)	(87.6)	(77.9)
...older than 2003	-29.2	-314.9***	-29.7	-86.3	-177.9*	20.6	-67.0	96.5
	(74.9)	(70.1)	(76.3)	(70.8)	(76.9)	(69.1)	(76.8)	(68.9)
...older than 2002	-417.2***	-202.4*	-468.8***	-236.0**	-399.4***	-297.1***	-568.3***	-418.8***
	(86.1)	(84.6)	(87.6)	(84.3)	(88.4)	(82.4)	(88.2)	(83.3)
...older than 2001	-379.4***	83.6	-389.8***	-542.5***	-445.7***	-340.2***	-402.9***	-308.9***
	(91.4)	(91.4)	(98.0)	(97.3)	(94.1)	(91.9)	(93.2)	(91.7)
5th-order age polynomial	X	X	X	X	X	X	X	X
Controls for car features	X	X	X	X	X	X	X	X
Controls for model updates	D1	D1	D2	D2	D3	D3	D4	D4
Car model fixed effects		X		X		X		X
R-squared	0.8179	0.8650	0.8203	0.8621	0.8169	0.8587	0.8163	0.8583
N	50,872	50,872	50,872	50,872	50,872	50,872	50,872	50,872

Notes: Identification strategy is based on a regression discontinuity (RD) design. The "...older than xxxx" dummy variables indicate whether a car has crossed a given year threshold. The corresponding coefficients are the parameters of interest, as they measure the discontinuous difference in prices – conditional on the exact car age – when comparing cars that were registered in different years.

D1: Model generation dummy (one for each model generation) imposing no insecurity: We treat all cars registered 1 month after official model switch as new model.
D2: Model generation dummy (one for each model generation) with 5-months insecurity windows: Like D1 but we treat model status of cars registered within the three months after a model switch as "unknown". Effectively, these cars are not used to identify our model. **D3:** Model update dummy, imposing no insecurity: Takes on the value of 1 if a model update happened in a given month. **D4:** Model update dummy, with 5-months insecurity windows: Like D3 but the three months after the introduction are also labeled as a model update month. **Main Specification:** We classify a car as having undergone a model update if its date of first registration was more than 3 months after factories switched production. Robust standard errors in brackets. ***p<0.001; **p<0.01; *p<0.05

Table 7. *Robustness Analysis Vintage: Varying age polynomials*

Dep. Variable: Car price	(1)	(2)	(3)	(4)	(5)
Indicator for...					
...older than 2007	-136.7	45.1	-325.7**	-606.1***	-677.9***
	(93.8)	(105.9)	(116.6)	(130.5)	(141.1)
...older than 2006	-276.8***	-654.7***	-575.8***	-436.3***	-397.0***
	(51.1)	(60.4)	(64.5)	(66.6)	(70.9)
...older than 2005	-512.7***	-908.7***	-306.0***	-330.0***	-389.4***
	(60.2)	(63.8)	(69.8)	(69.2)	(72.0)
...older than 2004	-898.5***	-514.2***	-205.7*	-499.4***	-465.3***
	(76.2)	(72.6)	(86.6)	(97.9)	(96.4)
...older than 2003	-266.7***	-199.9***	-464.8***	-334.4***	-269.0**
	(62.8)	(75.5)	(87.1)	(87.1)	(92.0)
...older than 2002	-160.6	-100.6	-319.2***	-119.8	-203.7*
	(68.3)	(81.3)	(88.1)	(93.4)	(96.7)
...older than 2001	114.2	-208.7*	-344.5***	-23.2	-272.5*
	(91.1)	(89.9)	(102.0)	(112.1)	(116.7)
Order of age polynomial	3rd	4th	6th	7th	8th
Controls for car features	X	X	X	X	X
Controls for model updates	X	X	X	X	X
Car model fixed effects	X	X	X	X	X
R-squared	0.8594	0.8185	0.8178	0.8179	0.8179
N	50,872	50,872	50,872	50,872	50,872

Notes: Identification strategy is based on a regression discontinuity (RD) design. The "...older than xxxx" dummy variables indicate whether a car has crossed a given year threshold. The corresponding coefficients are the parameters of interest, as they measure the discontinuous difference in prices – conditional on the exact car age – when comparing cars that were registered in different years. Robust standard errors in brackets. ***p<0.001; **p<0.01; *p<0.05

Table 8. *The impact of mileage discontinuities on price*

Dep. Variable: Car price	(1)	(2)	(3)	(4)	(5)
Indicator for...					
...more than 10K km	-1,582.4***	-994.5***	-1,022.9***	-888.7***	-905.9***
	(249.1)	(174.2)	(173.4)	(160.8)	(160.5)
...more than 20K km	-1075.1***	-789.7***	-779.4***	-758.0***	-749.2***
	(164.7)	(109.1)	(108.3)	(96.2)	(95.7)
...more than 30K km	-420.1**	-359.1***	-368.6***	-362.3***	-368.8***
	(133.8)	(83.1)	(82.7)	(71.6)	(71.4)
...more than 40K km	-576.4***	-372.6***	-378.8***	-424.1***	-428.4***
	(120.9)	(69.8)	(69.6)	(59.2)	(59.2)
...more than 50K km	-442.4***	-146.1*	-141.0*	-168.5***	-163.9***
	(113.7)	(60.5)	(60.5)	(51.6)	(51.6)
...more than 60K km	-101.1	-210.7***	-212.3***	-241.2***	-242.9***
	(111.3)	(57.5)	(57.6)	(48.9)	(48.9)
...more than 70K km	-586.8***	-221.3***	-220.0***	-235.9***	-234.6***
	(106.3)	(55.2)	(55.3)	(46.9)	(46.8)
...more than 80K km	-108.8	-143.9**	-143.8**	-213.3***	-212.9***
	(103.0)	(52.9)	(52.9)	(45.3)	(45.2)
...more than 90K km	-414.1***	-289.5***	-284.9***	-284.9***	-281.8***
	(100.4)	(50.7)	(50.7)	(43.8)	(43.8)
...more than 100K km	-71.2	-125.3*	-121.6*	-172.7***	-169.6***
	(103.5)	(53.1)	(53.1)	(46.4)	(46.4)
2nd-order mileage polynomial	X	X	X	X	X
Controls for car features		X	X	X	X
Controls for model updates			X		X
Car model fixed effects				X	X
R-squared	0.2634	0.8141	0.8148	0.8572	0.8575

Notes: Identification strategy is based on a regression discontinuity (RD) design. The "...more than xxK km" dummy variables indicate whether a car has crossed a given 10,000km threshold. The corresponding coefficients are the parameters of interest, as they measure the discontinuous difference in prices – conditional on the exact mileage – when comparing cars that have different mileage. Robust standard errors in brackets. ***p<0.001; **p<0.01; *p<0.05

Table 9. *Robustness Analysis Mileage: Placebo threshold US miles*

Dep. Variable: Car price	(1)
Indicator for...	
...more than 10K miles	-709.0***
	(176.7)
...more than 20K miles	-97.1
	(131.3)
...more than 30K miles	1.6
	(107.9)
...more than 40K miles	-181.0*
	(79.3)
...more than 50K miles	-105.6
	(161.2)
...more than 60K miles	-71.4
	(74.3)
...more than 70K miles	-45.3
	(86.3)
...more than 80K miles	34.9
	(126.8)
...more than 90K miles	-101.6
	(100.1)
...more than 100K miles	-35.4
	(85.1)
2nd-order mileage polynomial	X
Controls for car features	X
Controls for model updates	X
Car model fixed effects	X
R-squared	0.8172
N	50,872

Notes: Identification strategy is based on a regression discontinuity (RD) design. The "...more than xxK miles" dummy variables indicate whether a car has crossed a given 10,000mile threshold. The corresponding coefficients are the parameters of interest, as they measure the discontinuous difference in prices – conditional on the exact mileage – when comparing cars that have different mileage. All regressions also include the original non-placebo 10,000Km thresholds, for which estimates do not significantly change. Robust standard errors in brackets. ***p<0.001; **p<0.01; *p<0.05

Table 10. *Robustness Analysis Mileage: Log-linearization*

Dep. Variable: ln(Car price)	(1)	(2)	(3)	(4)	(5)
Indicator for...					
...more than 10K km	-0.057*	0.044***	0.041***	0.060***	0.058***
	(0.011)	(0.008)	(0.006)	(0.005)	(0.005)
...more than 20K km	-0.033***	0.029***	0.029***	0.034***	0.035***
	(0.009)	(0.006)	(0.004)	(0.003)	(0.004)
...more than 30K km	-0.001	0.050***	0.050***	0.052***	0.051***
	(0.008)	(0.005)	(0.005)	(0.004)	(0.004)
...more than 40K km	-0.015*	0.034***	0.033**	0.031***	0.030***
	(0.008)	(0.005)	(0.007)	(0.006)	(0.006)
...more than 50K km	-0.016*	0.007	0.008	0.005	0.006
	(0.008)	(0.004)	(0.008)	(0.006)	(0.006)
...more than 60K km	0.0107	0.006	0.006	0.003	0.003
	(0.008)	(0.004)	(0.009)	(0.008)	(0.008)
...more than 70K km	-0.030***	-0.004	-0.004	-0.006	-0.006
	(0.008)	(0.004)	(0.009)	(0.008)	(0.008)
...more than 80K km	0.006	0.000	0.000	-0.005	-0.005*
	(0.008)	(0.004)	(0.009)	(0.008)	(0.008)
...more than 90K km	-0.024***	-0.017***	-0.017***	-0.017***	-0.017***
	(0.008)	(0.004)	(0.009)	(0.008)	(0.008)
...more than 100K km	0.006	-0.002	-0.002	-0.005	-0.005
	(0.009)	(0.005)	(0.010)	(0.008)	(0.008)
5th-order age polynomial	X	X	X	X	X
Controls for car features		X	X	X	X
Controls for model updates			X		X
Car model fixed effects				X	X
R-squared	0.2930	0.8208	0.8215	0.8692	0.8699
N	63,340	50,872	50,872	50,872	50,872

Notes: Identification strategy is based on a regression discontinuity (RD) design. The "...more than xxK km" dummy variables indicate whether a car has crossed a given 10,000km threshold. The corresponding coefficients are the parameters of interest, as they measure the discontinuous difference in prices – conditional on the exact mileage – when comparing cars that have different mileage. Robust standard errors in brackets. ***p<0.001; **p<0.01; *p<0.05

Table 11. *Robustness Analysis Mileage: Varying controls for model updates*

Dep. Variable: ln(Car price)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Indicator for...								
...more than 10K km	-937.4***	-1,143.9***	-981.6***	-1,024.6***	-1,048.4***	-926.0***	-1,024.7***	-903.8***
	(173.9)	(157.2)	(173.5)	(158.0)	(170.7)	(158.9)	(173.6)	(160.7)
...more than 20K km	-762.3***	-788.6***	-775.0***	-756.6***	-739.2***	-722.1***	-784.0***	-754.0***
	(109.0)	(92.9)	(109.2)	(94.1)	(105.4)	(93.9)	(108.4)	(95.9)
...more than 30K km	-322.2***	-455.7***	-355.0***	-378.9***	-353.5***	-356.8***	-369.9***	-367.9***
	(82.9)	(69.9)	(83.5)	(70.7)	(81.4)	(70.5)	(82.9)	(71.5)
...more than 40K km	-351.3***	-480.9***	-374.2***	-442.3***	-367.3***	-419.2***	-380.9***	-428.5***
	(69.5)	(58.0)	(70.6)	(59.0)	(69.1)	(58.7)	(69.7)	(59.2)
...more than 50K km	-131.0*	-153.3**	-160.1**	-144.7**	-145.5*	-167.1**	-146.9*	-168.3**
	(60.3)	(50.4)	(61.5)	(51.7)	(60.6)	(51.7)	(60.5)	(51.6)
...more than 60K km	-208.1***	-214.4***	-226.7***	-214.3***	-203.7***	-235.2***	-209.1***	-240.3***
	(57.3)	(47.6)	(58.5)	(49.0)	(57.8)	(49.1)	(57.6)	(48.9)
...more than 70K km	-220.3***	-218.1***	-271.5***	-229.3***	-223.1***	-236.6***	-224.8***	-237.6***
	(54.9)	(45.3)	(56.2)	(47.0)	(55.5)	(47.0)	(55.2)	(46.8)
...more than 80K km	-144.0**	-201.8***	-134.4*	-187.5***	-142.1**	-211.1***	-138.8**	-209.8***
	(52.5)	(43.7)	(53.9)	(45.5)	(53.1)	(45.4)	(52.9)	(45.2)
...more than 90K km	-285.6***	-257.3***	-292.0***	-249.9***	-284.5***	-281.3***	-285.3***	-282.6***
	(50.2)	(42.8)	(51.6)	(44.5)	(50.9)	(44.0)	(50.7)	(43.8)
...more than 100K km	-122.9*	-182.4***	-146.0**	-165.9***	-122.9*	-169.9***	-128.5*	-174.2***
	(52.5)	(45.4)	(54.5)	(47.6)	(53.3)	(46.5)	(53.0)	(46.4)
2nd-order mileage polynomial	X	X	X	X	X	X	X	X
Controls for car features	X	X	X	X	X	X	X	X
Controls for model updates	D1	D1	D2	D2	D3	D3	D4	D4
Car model fixed effects		X		X		X		X
R-squared	0.8160	0.8651	0.8173	0.8623	0.8160	0.8581	0.8148	0.8574
N	50,872	50,872	48,572	48,572	50,872	50,872	50,872	50,872

Notes: Identification strategy is based on a regression discontinuity (RD) design. The "...more than xxK km" dummy variables indicate whether a car has crossed a given 10,000km threshold. The corresponding coefficients are the parameters of interest, as they measure the discontinuous difference in prices – conditional on the exact mileage – when comparing cars that have different mileage.
D1: Model generation dummy (one for each model generation) imposing no insecurity: We treat all cars registered 1 month after official model switch as new model.
D2: Model generation dummy (one for each model generation) with 5-months insecurity windows: Like D1 but we treat model status of cars registered within the three months after a model switch as "unknown". Effectively, these cars are not used to identify our model. **D3:** Model update dummy, imposing no insecurity: Takes on the value of 1 if a model update happened in a given month. **D4:** Model update dummy, with 5-months insecurity windows: Like D3 but the three months after the introduction are also labeled as a model update month. **Main Specification:** We classify a car as having undergone a model update if its date of first registration was more than 3 months after factories switched production. Robust standard errors in brackets. ***p<0.001; **p<0.01; *p<0.05

Table 12. *Robustness Analysis Mileage: Varying mileage polynomials*

Dep. Variable: Car price	(1)	(2)	(3)	(4)	(5)
Indicator for...					
...more than 10K km	-1,039.5***	-804.1***	-585.7***	-476.5**	999.2**
	(161.6)	(178.0)	(178.0)	(183.3)	(313.4)
...more than 20K km	-864.1***	-558.2***	-337.6***	-233.1	392.6*
	(98.9)	(117.6)	(118.6)	(126.9)	(184.4)
...more than 30K km	-463.1***	-143.2	67.9	161.7	214.3
	(75.3)	(93.6)	(93.8)	(101.9)	(127.2)
...more than 40K km	-524.7***	-175.4*	19.7	100.2	-151.8
	(62.6)	(81.7)	(82.5)	(90.2)	(115.0)
...more than 50K km	-249.5***	83.6	254.4***	-319.2***	-20.3
	(55.5)	(71.6)	(71.9)	(77.8)	(96.0)
...more than 60K km	-369.1***	-38.7	114.7	167.7*	-138.7
	(55.3)	(67.7)	(67.8)	(72.3)	(84.0)
...more than 70K km	-352.6***	-39.0	103.9	148.0*	-99.2
	(53.3)	(65.0)	(64.9)	(68.4)	(78.6)
...more than 80K km	-344.9***	10.7	131.4 *	163.8**	6.4
	(51.9)	(60.7)	(60.4)	(62.6)	(75.5)
...more than 90K km	-412.2***	-126.2*	-15.9	9.0	-96.4
	(50.0)	(58.0)	(57.3)	(58.9)	(75.6)
...more than 100K	-297.5***	-11.6	77.6	93.7	30.8
	(51.2)	(58.2)	(57.7)	(58.5)	(74.9)
Order of mileage polynomial	1st	3rd	4th	5th	6th
Controls for car features	X	X	X	X	X
Controls for model updates	X	X	X	X	X
Car model fixed effects	X	X	X	X	X
R-squared	0.8605	0.8183	0.8184	0.8184	0.8178
N	50,872	50,872	50,872	50,872	50,872

Notes: Identification strategy is based on a regression discontinuity (RD) design. The "...more than xxK km" dummy variables indicate whether a car has crossed a given 10,000km threshold. The corresponding coefficients are the parameters of interest, as they measure the discontinuous difference in prices – conditional on the exact mileage – when comparing cars that have different mileage. Robust standard errors in brackets. ***p<0.001; **p<0.01; *p<0.05

Table 13. *Robustness Analysis: The joint impact of Vintage and Mileage discontinuities on price*

Dep. Variable: Car price	(1)	(2)	(3)
Indicator for...			
...older than 2007	-221.4*		-265.7**
	(91.8)		(90.1)
...older than 2006	-780.3***		-805.9***
	(56.6)		(55.5)
...older than 2005	-609.2***		-610.0***
	(54.9)		(54.1)
...older than 2004	-3.7		7.9
	(77.7)		(71.9)
...older than 2003	44.3		56.4
	(68.9)		(67.5)
...older than 2002	-437.9***		-441.3***
	(83.5)		(83.2)
...older than 2001	-345.1***		-385.8
	(91.7)		(90.6)
...more than 10K km		-905.9***	-719.3**
		(160.5)	(159.5)
...more than 20K km		-749.2***	-550.9***
		(95.7)	(93.4)
...more than 30K km		-368.8***	-140.5*
		(71.4)	(69.8)
...more than 40K km		-428.4***	-218.2***
		(59.2)	(56.9)
...more than 50K km		-163.9***	-84.2
		(51.6)	(49.9)
...more than 60K km		-242.9***	-174.2***
		(48.9)	(47.6)
...more than 70K km		-234.6***	-190.7***
		(46.8)	(45.7)
...more than 80K km		-212.9***	-210.8***
		(45.2)	(44.0)
...more than 90K km		-281.8***	-290.3***
		(43.8)	(42.8)
...more than 100K km		-169.6***	-185.0***
		(46.4)	(45.3)
5th-order age polynomial	X		X
2nd-order mileage polynomial		X	X
Controls for car features	X	X	X
Controls for model updates	X	X	X
Car model fixed effects	X	X	X
R-squared	0.8584	0.8575	0.8661
N	50,872	50,872	50,872

Notes: Identification strategy is based on a regression discontinuity (RD) design. The "...older than xxxx" and "...more than xxK km" dummy variables indicate whether a car has crossed a given year or 10,000km threshold. The corresponding coefficients are the parameters of interest, as they measure the discontinuous difference in prices when comparing cars that have different age and mileage. Robust standard errors in brackets. ***p<0.001; **p<0.01; *p<0.05