

# DISCUSSION PAPER SERIES

No. 3360

**THE MARKET FOR USED CARS:  
A NEW TEST OF THE LEMONS  
MODEL**

Winand Emons and George Sheldon

***INDUSTRIAL ORGANIZATION***



**Centre for Economic Policy Research**

**[www.cepr.org](http://www.cepr.org)**

Available online at:

[www.cepr.org/pubs/dps/DP3360.asp](http://www.cepr.org/pubs/dps/DP3360.asp)

# THE MARKET FOR USED CARS: A NEW TEST OF THE LEMONS MODEL

**Winand Emons**, Universität Bern and CEPR  
**George Sheldon**, Universität Basel

Discussion Paper No. 3360  
May 2002

Centre for Economic Policy Research  
90–98 Goswell Rd, London EC1V 7RR, UK  
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999  
Email: [cepr@cepr.org](mailto:cepr@cepr.org), Website: [www.cepr.org](http://www.cepr.org)

This Discussion Paper is issued under the auspices of the Centre's research programme in **INDUSTRIAL ORGANIZATION**. Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as a private educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions. Institutional (core) finance for the Centre has been provided through major grants from the Economic and Social Research Council, under which an ESRC Resource Centre operates within CEPR; the Esmée Fairbairn Charitable Trust; and the Bank of England. These organizations do not give prior review to the Centre's publications, nor do they necessarily endorse the views expressed therein.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

Copyright: Winand Emons and George Sheldon

May 2002

## ABSTRACT

### The Market for Used Cars: A New Test of the Lemons Model\*

The lemons model assumes that owners of used cars have an informational advantage over potential buyers with respect to the quality of their vehicles. Owners of bad cars will try to sell them to unsuspecting buyers while owners of good cars will hold on to theirs. Consequently, the quality of traded automobiles should be sub-average. In contrast to previous work, the following Paper tests both the assumption of informational asymmetry and the prediction of sub-average traded car quality using a sample consisting of all 1985 cars registered in the Swiss canton of Basle-City over the period 1985-91. Our data support both the assumption and the prediction of the lemons model. The lemons problem does not appear to be widespread, however.

JEL Classification: C41, D82, L15 and L62

Keywords: adverse selection, duration models and used car market

Winand Emons  
VWI  
Universität Bern  
Gesellschaftsstr. 49  
CH-3012 Bern  
SWITZERLAND  
Tel: (41 31) 631 3922  
Fax: (41 31) 631 3992  
Email: emons@vwi.unibe.ch

George Sheldon  
Labor & Industrial Economics  
Universität Basel  
Petersgraben 51  
CH-4003 Basel  
SWITZERLAND  
Tel: (41 61) 267 0311  
Fax: (41 61) 267 3035  
Email: sheldon@ubaclu.unibas.ch

For further Discussion Papers by this author see:  
[www.cepr.org/pubs/new-dps/dplist.asp?authorid=116192](http://www.cepr.org/pubs/new-dps/dplist.asp?authorid=116192)

For further Discussion Papers by this author see:  
[www.cepr.org/pubs/new-dps/dplist.asp?authorid=146640](http://www.cepr.org/pubs/new-dps/dplist.asp?authorid=146640)

\* Financial support from the WWZ-Forum at the University of Basle is gratefully acknowledged. We further thank the Department of Motor Vehicles of Basle-City (Switzerland) for supplying us with the data used in this study, as well as Tom Flury and Frank Losert for painstakingly transcribing the vehicle safety inspection records into electronic form. Thanks are also due to Andreas Diekmann, David Gensove, Kurt Schmidheiny and Pedro Pita Barros for helpful comments.

Submitted 15 April 2002

# 1. Introduction

In AKERLOF's (1970) "lemons" model, buyers cannot distinguish between the qualities of used cars. Consequently, cars of different qualities sell at a common price that reflects the buyers' estimate of average quality. Sellers, on the other hand, know the quality of their vehicles, yet due to the non-observability of car quality to buyers, can only sell their vehicles at the common price reflecting average quality. Hence, owners of automobiles of above-average quality hold on to their cars, while owners of "lemons" attempt to sell theirs to unsuspecting buyers. The average quality of traded cars is thus lower than the average quality of the whole population of cars: good used cars are driven out of the market by bad ones. In the limit this adverse selection phenomenon may even lead to market breakdown.

Empirical tests of the lemons model customarily concentrate on adverse selection. In line with this approach, the quality of traded goods is commonly compared with that of the population of goods from which they stem. The results of such comparisons often show that lower quality goods indeed trade with greater frequency. These findings are inconclusive, however, as buyers may be fully informed and nevertheless purchase lemons because they are more willing than sellers to accept the associated risks and/or costs of owning them. The following study extends previous work on the lemons problem by testing for both adverse selection and quality uncertainty. By showing that cars that are put on the market represent a negative selection out of the population of registered cars and that the buyers of these cars suffer from quality uncertainty, we establish that sellers enjoy an informational advantage over buyers in the used car market.

We apply our test of the lemons model on a sample consisting of all 1985 cars registered in the Swiss canton of Basle-City over the period 1985-1991. Our approach proceeds in two steps. First we demonstrate that car owners put a disproportionate share of sub-average quality cars up for sale, and then we show that the buyers of these cars suffer from quality uncertainty. Our test of adverse selection thus focuses on sellers, while our test of quality uncertainty concentrates on buyers.

Our test of adverse selection is based on the results of vehicle safety-inspection records. Cars in Basle-City are subject to mandatory safety inspections every three years. Under certain conditions, however, vehicles are required to be inspected more frequently if sold, thus enabling us to discriminate between the quality of sold and non-sold cars. Our results indicate that the probability of a vehicle having a defect in-

creases if the car was sold privately, confirming the presence of adverse selection. Interestingly, we find that the opposite holds true for used cars sold by dealers, an observation that conforms with AKERLOF's (1970) original conjecture that institutions such as warranties and reputation may develop to offset the effects of asymmetric information.<sup>1</sup>

Our test of quality uncertainty rests on the following simple idea. The "lemons" model assumes that buyers of used cars are incompletely informed about the true quality of the vehicles they purchase. With time, however, owners should discover the true condition of their purchased vehicles. Hence, those individuals who bought lemons should become increasingly aware of their mistakes and attempt to unload the lemons on other unassuming buyers. As a consequence, the probability of a recently acquired car being re-sold should, after perhaps initially increasing, decline as ever fewer defects remain to be discovered. To test this conjecture, we estimate the hazard function of car ownership, which - roughly speaking - gives the probability of selling one's car conditional on the length of current ownership. If car buyers suffer from quality uncertainty at the time of purchase, the hazard function should decline with the duration of ownership, reflecting the decreasing likelihood that buyers will discover a new defect and attempt to re-sell their vehicles.

The hazard function approach has been employed before in labor market studies to test for quality uncertainty, which also underlies job-matching models of labor turnover (JOVANOVIC, 1979).<sup>2</sup> The following study appears to be the first attempt to employ this methodology to the lemons model, however. Using the ownership biographies of all 1985 cars registered in Basle-City in the sample period, we show that the estimated hazard function of car ownership does indeed decline with the duration of ownership as quality uncertainty implies. In addition, we find that higher quality cars are less likely of being re-sold shortly after purchase.

---

<sup>1</sup> In the US all states have "lemon" laws, though they differ as to what they require dealers to do with returned vehicles. It is still possible to sell a car identified in one state as a lemon without warning in another state (see [http://www.autopedia.com/html/HotLinks\\_Lemon.html](http://www.autopedia.com/html/HotLinks_Lemon.html)). In the EU consumers enjoy less protection. A new directive taking effect in 2002 will force dealers to fix defects within two years in a new car or one year in a used car (see <http://europa.eu.int/scadplus/leg/en/lvb/l32022.htm>).

<sup>2</sup> Job-matching models assume that an employee and employer are not perfectly informed as to whether they meet each other's expectations when entering into an employment contract. In the course of time, however, "mismatches" are discovered and get sorted out by job separations. Hence, the hazard function of job tenure should, after perhaps initially rising, decrease thereafter. BORLAND and LYE (1996), CHAPMAN and SOUTHWICK (1991), HAUTSCH ET AL. (2001), LANCASTER ET AL. (1987), OHTAKE and OHKUSA (1994), and SHELDON (1991) find empirical support for this conjecture.

The remainder of the paper is organized as follows. The next section provides a short overview of previous empirical studies of the lemons problem. Section 3 describes the data upon which our study is based. Section 4 explains our method of testing for adverse selection more fully and submits our findings. Section 5 describes our test of quality uncertainty in depth and presents our results. The last section summarizes and concludes.

## 2. Previous Empirical Studies

Empirical studies on adverse selection can be divided into roughly two groups: those based on quality and those based on price comparisons. If the lemons model holds, then quality comparisons should show that the quality of traded goods is sub-average, while price comparisons should find that goods prone to adverse selection sell at lower prices than comparable goods not subject to negative selection.

Studies employing quality comparisons include the following.

- BOND (1982, 1984), like us, applies quality comparisons to the market for used motor vehicles. He compares the frequency of maintenance of pickup trucks purchased new and those purchased used. His results indicate that if the effects of age and total mileage are controlled for, there is no difference in maintenance for trucks acquired new and those acquired used and less than 10-years old; for older trucks, however, he finds evidence of more frequent repairs among those traded recently.<sup>3</sup>
- LEHN (1984) looks instead at major-league baseball. After six years of service, players can sign new contracts with their original team or become free agents and sign with new teams. LEHN finds evidence that free agents spend more time on the disabled list than renewing agents, supporting the adverse-selection hypothesis.
- GREENWALD and GLASSPIEGEL (1983) focus on the market for slaves in pre-Civil War New Orleans. According to the lemons model, better informed slave-owners are likely to sell off their poorest quality slaves to uninformed buyers. Indeed, they find that slaves brought to market were of 20-40% lower quality, on average, than the slave population in general, and that good slaves were three times less likely to be sold than low quality ones.

---

<sup>3</sup> BOND (1982) initially found no evidence of the lemons problem. PRATT and HOFFER (1984) challenged this result. In BOND's (1984) response he reports finding adverse selection for old trucks.

Other authors apply the price-comparison approach.

- GENESOVE (1993) compares the prices of used cars sold in the wholesale used-car market by used and new-car dealers. New-car dealers (selling both new and used cars) tend to sell a larger share of their trade-ins indiscriminately in the wholesale market than used car dealers (selling only used cars). Adverse selection suggests that new car dealers will sell cars of higher quality, on average, and receive a higher price in return. GENESOVE uses price data collected at a wholesale auction to test this prediction. He finds evidence for adverse selection.
- FABEL and LEHMANN (2000) argue that internet used-car markets give rise to lower search costs than do traditional used-car markets, whereas traditional markets are less prone to adverse selection. In equilibrium the two markets may co-exist: e-commerce trade occurs across all quality segments while the traditional markets are devoid of low-quality vehicles. Using price data from internet and print media advertisements, they find that average prices of cars traded electronically are indeed lower than those traded traditionally.
- CHEZUM and WIMMER (1997) study thoroughbred yearling auctions. They distinguish between sellers who breed horses to race and sell (racers) and sellers who take all their yearlings to auctions (breeders). If racers use private information, keeping those yearlings with a high probability of on-track success, they are likely to receive a lower price for similar yearlings as compared to breeders. The authors find support for this hypothesis.
- ROSENMAN and WILSON (1991) analyze cherry markets. There producers who do not sort their cherries must sell them in the market with the minimum size standard corresponding to the smallest cherries in the lot. If a firm wishes to sell its cherries according to a larger size standard, the cherries must be sorted. Since sorting is costly, a firm may decide not to sort and instead sell its higher quality cherries mixed with smaller cherries at a lower standard. Sorting firms will sell more lower quality cherries in the minimum size market than non-sorting firms do, and prices should reflect this. Indeed, prices are higher for non-sorted cherries.
- GIBBONS and KATZ (1991) compare subsequent wages of workers displaced by plant closing and those laid off. When firms have discretion as to whom to lay off, the market infers that laid-off workers are of lower ability. No such negative inference is warranted if workers are displaced by a plant closing. Accordingly, post-displacement wages should be lower and post-displacement unemployment spells should be longer for those displaced by layoffs than for those displaced by plant closings, and pre-displacement wages should not differ by cause of displacement.

The evidence supports all three predictions.<sup>4</sup>

### 3. Data

Our study is based two separate sets of data. The one consists of protocols from state vehicle safety inspections, and the other of registration records drawn from the data-bank of the Department of Motor Vehicles in Basle-City. All data pertain to 1985 cars registered in Basle-City at some time during the sample period from 1985 through 1991.

The vehicle safety inspection protocols constitute form sheets filled out by the state inspector on the day the car was initially presented for inspection. The inspection protocols document a vehicle's chassis number (unique to each vehicle), the initial date of inspection, vehicle mileage, the discovered defects, the make of the car and the grounds for presenting the vehicle. Six different reasons for submitting an automobile for inspection are recorded: (i) periodic<sup>5</sup> check up, (ii) private or (iii) dealer sale<sup>6</sup>, (iv) ticketing by police for a safety defect, (v) technical altering, and (vi) export out of Switzerland. In all instances, an inspection is mandatory and not left to the choice of an owner.

The owner of cars that do not initially pass inspection must have the defects repaired and re-submit their vehicles until they can pass inspection. The dates of passed inspections are documented in the registration papers belonging to the vehicle. These remain with the vehicle throughout its life and are passed on to the next owner at every new sale. Hence, a buyer of a used car knows when the vehicle was last inspected and can be assured that all discovered safety defects were repaired.

We use these protocol data to investigate whether cars put up for sale are more or less likely to display a defect than a randomly chosen vehicle (*Section 4*). The latter is de-

---

<sup>4</sup> Several papers deal with the presence of adverse selection in insurance markets. CUTLER and ZECKHAUSER (2000) conclude that adverse selection is large in health insurance markets. PUELZ and SNOW (1994) present empirical evidence of adverse selection in the market for automobile collision insurance. DIONNE ET AL. (2001) challenge this result and argue that when non-linearity is taken into account, adverse selection vanishes. CHIAPPORI and SALANIÉ (2000) find no adverse selection using data on inexperienced French drivers. COHEN (2001) finds adverse selection using data on experienced Israeli drivers.

<sup>5</sup> Due to a lack of personnel, most cars were inspected at four-year intervals in the sample period although three-year intervals were mandatory. Vehicles used for commercial purposes (taxis, rental, driver education) have to appear for inspection annually.

<sup>6</sup> Cars less than four-years old at the time of sale or that had been inspected in the last year prior to sale were not subject to inspection in the sample period.

defined as a car submitted for a periodic inspection [reason (i) given above]. The unit of observation in this analysis is an inspection protocol.

The registration records, as the name implies, document all vehicle registrations, de-registrations and re-registrations that occurred in Basle-City in the sample period. From these records we construct 3,806 spells of vehicle ownership, 2,645 of which pertain to cars bought new and 1,161 to automobiles purchased used. An ownership spell begins with the registration of a car by the current owner and ends with the final de-registration of the vehicle by the same individual. A spell is viewed as being complete when a new owner re-registers the same car, which implies that a sale has occurred.<sup>7</sup> Otherwise the spell is considered to be incomplete, or right-censored, and treated appropriately in estimation.

The spells document the chassis number of a vehicle, its make, engine type (diesel or gasoline), beginning and end date of a spell, the current owner's social security number, the license plate number, the vehicle's current use (private or commercial; taxi, rental, or driver education), and whether the spell ended with a sale. We use the ownership spells to analyze the shape of the hazard function, i.e., to investigate how resale probabilities evolve as the length of vehicle ownership increases (*Section 5*). The unit of observation in this analysis is an ownership spell.

The possibilities of merging the two sets of data are limited. For example, the information on vehicle mileage contained in the inspection protocols is far too spotty to be integrated into the spell data. The only protocol information pertinent to the hazard function analysis is whether a vehicle was inspected prior to being sold, which can be determined by matching inspection dates with spell dates. This we did. On the other hand, whether a vehicle initially passed inspection or not is of little relevance since any discovered defects were repaired, a fact known both to buyer and seller. It is the undiscovered defects that lead to quality uncertainty.

## 4. Test of Adverse Selection

Our test of adverse selection is based on a comparison of the quality of sold and non-sold cars, where the definition of car quality rests on the results of official vehicle

---

<sup>7</sup> The protocol records cannot be used to define a sale since not all cars sold are subject to inspection. Moreover, a car brought in for a mandatory periodic check could nevertheless be sold afterwards. Note that a car must be sold within the confines of Basle-City for a sale to be observed with our registration data.

safety inspections.<sup>8</sup> We consider cars that passed inspection at their initial presentation to be of higher quality than those that failed. We use a probit model based on the following threshold model to investigate whether cars put up for sale are more likely to fail an inspection test.

Let

$$y_i^* = -\boldsymbol{\beta}'\mathbf{x}_i + \varepsilon_i, \quad (1)$$

where  $y_i^*$  = quality of vehicle  $i$  germane to passing inspection,  
 $\mathbf{x}_i$  = vector of negative covariates of quality,<sup>9</sup>  
 $-\boldsymbol{\beta}'\mathbf{x}_i$  = mean quality of vehicles with the observed profile of characteristics  $\mathbf{x}_i$ ,  
 $\varepsilon_i$  = normal i.i.d. random variable with expected value 0 and variance  $\sigma^2$ .

In addition, let  $\gamma$  be a threshold level of quality sufficient to pass inspection, and  $y_i$  be an indicator variable denoting whether a car passed inspection ( $y_i = 0$ ) or not ( $y_i = 1$ ). Then clearly  $y_i = 1$  if  $y_i^* < \gamma$ , and  $y_i = 0$  otherwise. Hence, the probability  $P(y_i = 1 | \mathbf{x}_i)$  that a car with observed profile  $\mathbf{x}_i$  fails inspection equals

$$\begin{aligned} P(y_i = 1 | \mathbf{x}_i) &= P(-\boldsymbol{\beta}'\mathbf{x}_i + \varepsilon_i < \gamma) \\ &= \Phi(\boldsymbol{\alpha}'\mathbf{x}_i), \end{aligned} \quad (2)$$

where  $\Phi$  denotes the standard normal distribution and  $\alpha_0 = (\gamma + \beta_0)/\sigma$  and  $\alpha_1, \dots, \alpha_K = \beta_1/\sigma, \dots, \beta_K/\sigma$ . Our specification thus implies that the covariates influence solely the mean quality of cars of a given profile  $\mathbf{x}_i$  and not the variation about the mean.

The covariate vector  $\mathbf{x}$  contains (i) variables which describe the car and (ii) dummy variables which distinguish between the various grounds for submitting a vehicle for inspection. We measure all descriptive variables in deviations from their respective means and choose vehicles brought in for a periodic inspection as the reference group. Hence, the coefficients of the dummy variables measure the relative chances of a non-periodically inspected car with average attributes passing inspection compared to an identical randomly chosen vehicle.

---

<sup>8</sup> EMONS and VON HAGEN (1991) use similar safety inspection data to establish weak information efficiency in the German automobile market.

<sup>9</sup> Vectors appear in bold type throughout.

Our sample encompasses 5,333 inspections conducted on 4,418 cars. *Table 1* describes the data contained in the inspection protocols.<sup>10</sup> The first column lists the various possible grounds for submitting a car for inspection. "Periodic checks" refers to the regular inspections to which cars in Basle-City are subject at fixed intervals. The table indicates that the large majority (63 percent) of cars submitted for inspection came in for a periodic check. The second largest group (30 percent) consisted of cars put up for sale or sold by dealers ("dealer sale").

The next columns give the share of inspected cars with diesel engines and the average mileage of the vehicles at the time of inspection measured in kilometers. The last columns in the table pertain to types of defects for which the safety inspections check. "Devices" refer to various minor items like windshield wipers and rearview mirrors. As the table indicates, just over 77 percent of the cars inspected exhibited no defects. The most common faults were linked to lights or the power train (engine and drive shaft). The table also shows that a below-average share (52.7 percent) of privately sold cars and an above-average share (83.3 percent) of dealer-sold vehicles had no technical defects, suggesting that cars offered for sale by private individuals are of lower quality than those vehicles dealers offer. However, this observation may simply reflect the fact that cars privately sold generally have more mileage on them. The probit model serves to control for this and other observed differences.

The maximum likelihood estimates of the parameter vector  $\alpha$  of the probit model appear in *Table 2*.<sup>11</sup> As can be seen, greater mileage does indeed raise the chances that a car will have a defect. Nonetheless, privately sold cars still show an above-average propensity of being defective. In every category, save one ("handbrake"), a privately sold car that went through a safety inspection was more likely to have a defect than a randomly chosen car with the same characteristics.<sup>12</sup> The opposite holds true for dealer-sold cars, which is in line with AKERLOF's (1970) conjecture that institutions may develop to offset the effects of asymmetric information. Based on the results appearing in the final column ("None"), a privately sold car has a 27 percent higher and a dealer-submitted car a 10 percent lower chance of having a defect than a randomly selected vehicle. Cars ticketed by the police for safety reasons ("ticketed") generally have an even lower likelihood of failing inspection. This result may seem surprising at

---

<sup>10</sup> The brand of car, also included in our analysis, is excluded from *Table 1* to preserve space. Our sample contains 45 different makes of car.

<sup>11</sup> The estimated coefficients of the dummy variables for vehicle brands are omitted from the table to preserve space.

<sup>12</sup> If one ignores the seven cars inspected for export and the four for which the grounds for inspection are missing, the reference group (regression constant) pertains to an average car called in for a periodic safety inspection.

first, but is probably due to the fact that the owner has the ticketed defect repaired before submitting the car for inspection. That is probably also the reason why cars sold by dealers are more likely to pass inspection: they check out their vehicles beforehand and repair anything they find wrong.<sup>13</sup>

The estimated regression constants represent z-scores for a randomly chosen vehicle. Thus the lower the value of a constant the less likely a particular type of defect will be found. A comparison of these estimated z-scores shows that randomly chosen cars are least likely to exhibit a defect related to steering or brakes, aspects of a car with which an owner through driving should be very familiar. On the other hand, defects pertaining to the frame and power train, which probably only an expert can discover, are more likely to turn up. The fact that lights are more likely to exhibit defects may reflect the fact that faulty lights only blind the oncoming traffic and go mostly unnoticed to the driver.

More important for the further analysis, however, is the finding that cars sold privately are more likely to have a defect than a randomly chosen vehicle. This clearly suggests that sellers are generally aware of the sub-average quality of the cars they sell. In the next section we investigate how buyers react to the cars they purchase.

---

<sup>13</sup> It is of course possible that the inspection results are biased in favor of dealers because inspectors trust dealers more and hence inspect their cars less thoroughly. The results in the next section based on buyer behavior suggest, however, that cars presented for inspection by dealers are also of higher quality with regard to unobserved characteristics.

*Table 1: Sample Means of Vehicle Inspections*

<b>Grounds</b>	Cases	Diesel	Kilometers	<b>Type of Defect:</b>								
				Lights	Devices	Frame	Suspension	Steering	Power Train	Brakes	Handbrake	None
Periodic Check	3363	0.050	61020	0.132	0.077	0.051	0.047	0.026	0.115	0.023	0.007	0.751
Private Sale	165	0.133	66924	0.248	0.255	0.152	0.103	0.042	0.218	0.042	0.006	0.527
Dealer Sale	1626	0.018	51954	0.073	0.033	0.021	0.037	0.020	0.077	0.017	0.005	0.833
Ticketed	54	0.093	61752	0.056	0.000	0.056	0.056	0.019	0.037	0.019	0.000	0.852
Vehicle Altered	114	0.123	49727	0.053	0.070	0.079	0.079	0.018	0.026	0.009	0.000	0.789
Export	7	0.143	78896	0.000	0.000	0.000	0.000	0.000	0.143	0.000	0.000	0.857
Unknown	4	0.250	80071	0.000	0.250	0.000	0.000	0.250	0.000	0.000	0.000	0.500
Total	5333	0.045	58242	0.115	0.068	0.046	0.046	0.024	0.104	0.021	0.006	0.771

Table 2: Probit Estimates of the Probability of a Vehicle Defect

Explanatory Variable	Type of Defect:								
	Lights	Devices	Frame	Suspension	Steering	Power Train	Brakes	Handbrake	None
<b>Constant</b>	-1.163*** (0.029)	-1.473*** (0.033)	-1.690*** (0.039)	-1.729*** (0.040)	-2.036*** (0.051)	-1.294*** (0.031)	-2.127*** (0.057)	-2.601*** (0.090)	0.715*** (0.024)
<b>Private Sale</b>	0.404*** (0.112)	0.743*** (0.113)	0.577*** (0.127)	0.404*** (0.141)	0.171 (0.187)	0.391*** (0.117)	0.275 (0.188)	-0.052 (0.380)	-0.563*** (0.102)
<b>Dealer Sale</b>	-0.342*** (0.057)	-0.407*** (0.071)	-0.408*** (0.083)	-0.109 (0.072)	-0.094 (0.091)	-0.189*** (0.057)	-0.099 (0.096)	-0.115 (0.154)	0.270*** (0.046)
<b>Ticketed</b>	-0.566** (0.290)	--	-0.028 (0.289)	-0.014 (0.294)	-0.251 (0.437)	-0.637** (0.325)	-0.242 (0.438)	--	0.484** (0.219)
<b>Vehicle Altered</b>	-0.460** (0.198)	0.024 (0.006)	0.270 (0.179)	0.307* (0.180)	-0.192 (0.306)	-0.671*** (0.253)	-0.316 (0.387)	--	0.076 (0.136)
<b>Mileage (km x 10<sup>-5</sup>)</b>	0.356*** (0.052)	2.320*** (1.231)	0.194** (0.070)	0.275*** (0.066)	0.419*** (0.071)	0.470*** (0.053)	0.262*** (0.086)	5.286*** (3.524)	-0.463*** (0.047)
<b>Diesel</b>	-0.017 (0.116)	-0.202* (0.185)	0.058 (0.144)	-0.213 (0.160)	0.177 (0.176)	-0.158 (0.128)	0.270 (0.186)	0.597 (0.316)	-0.104 (0.097)
<b>Sample Size</b>	5333	5333	5333	5333	5333	5333	5333	5333	5333
<b>lnL(<math>\beta_0</math>)</b>	-1900.40	-1323.22	-990.95	-997.02	-611.24	-1778.71	-547.33	-200.71	-2871.62
<b>lnL(<math>\beta^*</math>)</b>	-1790.83	-1240.03	-926.34	-956.57	-571.84	-1645.16	-506.14	-185.72	-2694.48
<b>-2[lnL(<math>\beta_0</math>)-lnL(<math>\beta^*</math>)]</b>	219.15***	166.37***	129.22***	80.90***	78.79***	267.09***	82.38***	29.97*	354.26***
<b>LRI [1-lnL(<math>\beta^*</math>)/lnL(<math>\beta_0</math>)]</b>	0.058	0.063	0.065	0.041	0.064	0.075	0.075	0.075	0.062

Asterisks denote statistical significance at the 10% (\*), 5% (\*\*), or 1% (\*\*\*) level. Estimated standard errors appear in parentheses.

## 5. Test of Quality Uncertainty

Our test of quality uncertainty among car buyers is based on the shape of the hazard function  $h(t)$  of car ownership. Roughly speaking, a hazard function gives the probability of an event (here car re-sale) occurring at time  $t$ , given that it has not occurred up to this point.<sup>14</sup> If car buyers suffer from quality uncertainty, the first months of ownership will consist of a learning process in which buyers discover the hidden characteristics of their newly purchased vehicles. Assuming that at least some buyers react to any dissatisfaction with their purchases by re-selling their vehicles, we should observe that the probability of a newly purchased car being re-sold, after possibly increasing during an initial phase of vehicle testing, declines with the duration of ownership, reflecting the falling number of defects remaining to be discovered. If, on the other hand, buyers do not suffer from quality uncertainty at the time of purchase, then any observed re-sales should merely stem from changes in buyers' tastes or constraints, which does not imply any particular pattern of duration dependence for the hazard function.<sup>15</sup>

To be more precise, let  $\bar{\mu}$  be the consumer's belief about the average expected quality of traded cars;  $\bar{\mu}$  is of course the actual expected quality so that expectations are correct. The actual quality of a vehicle at some point in time is the sum of two components. The first is the actual expected quality  $\mu$ , unknown to the buyer, which is a random draw from the distribution of all traded used cars. The second is a purely random component,  $\varepsilon_t$ , varying over time independently of  $\mu$ .

---

<sup>14</sup> Strictly speaking, the hazard function  $h(t)$  - like the corresponding probability density function  $f(t)$  - exhibits zero probability mass for given  $t$ . Hence the probability  $p(t)$  of selling one's car in a given interval  $t$  to  $t+dt$ , given that one has not done so by time  $t$ , really equals  $h(t) \cdot dt$ , where  $dt$  represents a small increase in time. Dividing the conditional probability  $p(t)$  by  $dt$  and letting  $dt$  go to zero yields the hazard function  $h(t)$ . The hazard function can also be viewed as the conditional density  $f(t)/S(t)$ , where  $S(t)$ , the so-called survivor function, gives the probability of a spell lasting at least to time  $t$ , i.e.,  $S(t) = 1 - F(t)$ , where  $F(t)$  is the cumulative probability distribution to  $f(t)$ . In the following, we also term  $p(t)$  the hazard rate. KALBFLEISCH and PRENTICE (1980) or LANCASTER (1990) provide extensive discussions of duration models, as hazard function models are also known. KIEFER (1988) gives a concise overview of the various approaches.

<sup>15</sup> Inasmuch as tastes towards a vehicle are more likely to change the longer one owns it, as at some point everyone begins to tire of his or her car, one might expect the hazard rate of re-sale to rise with the duration of ownership. This time pattern is the diametrical opposite of what quality uncertainty predicts, however. BULOW (1982) gives baby carriages as an example. In 1954 his parents had a high demand for one unit. In the current period their demand is zero, while someone else is doing the demanding. PORTER and SATTLER (1999) consider the following motive: The quality of a durable deteriorates over time so that current owners sell to update to their preferred quality. Their model predicts that vehicles depreciating relatively quickly are traded relatively frequently which is confirmed by their estimations. However, unlike our findings the trading pattern of new and used car buyers is similar. HENDEL and LIZZERI (1999) consider a dynamic model of adverse selection to examine the interactions of new and used car markets. Their model predicts that unreliable brands have steeper price declines and lower volumes of trade. This prediction is neither confirmed by their, by PORTER and SATTLER's, nor by our data; see *Table 3*.

Observation of the actual quality over time provides information about  $\mu$  so that at each subsequent date there exists a posterior distribution for the value of  $\mu$  given the stream of services the vehicle has provided to date. Let  $\eta$  be the mean of this distribution which is a function of the total services delivered by the car and the time that has elapsed.  $\eta$  is the expected value of  $\mu$  given the services the vehicle has provided. The value of  $\eta$  at the time of purchase is the prior mean of  $\mu$ , namely  $\bar{\mu}$ .

There will be values of  $\eta$  sufficiently low such that it is reasonable for the buyer to sell the vehicle and buy another one where the initial quality is again  $\bar{\mu}$ . The value of  $\eta$  below which the consumer sells will depend on time, since the precision of the posterior distribution of  $\mu$  increases with time. A low value of  $\eta$  early on will be discounted because little information about  $\mu$  has been accumulated and its value is, as yet, pretty uncertain. The same value of  $\mu$  later on when  $\mu$  is known rather accurately could well lead to a sale. Accordingly, we can imagine a threshold function of time  $\phi(t)$ , starting initially at some value below  $\bar{\mu}$ , such that the car is sold at the first instance (if it ever happens) at which the value of  $\eta$  falls below the value of the threshold function  $\phi$ . The threshold function, of course, reflects the transaction cost of selling the vehicle and buying another one.

This model has the following implications. On the one hand, the probability of a sale will initially be rather small because little learning has taken place and there is insufficient evidence on which to base a rational selling decision. On the other hand, eventually  $\mu$  is known with certainty, learning has ceased, and if a sale has not taken place by then, it never will (ignoring other motives to sell a car). Consequently, the probability of sale should initially be zero and eventually fall to zero.

From the standpoint of the test, it is not important whether a purchased car was actually defect or whether the buyer simply discovered that the car did not meet up to expectations or really fit his or her unchanged needs. In both cases quality uncertainty still holds. Our test does, however, require that at least some buyers express any unexpected dissatisfaction by re-selling their cars. Since the transaction costs of re-selling a newly purchased vehicle may be high, buyers could be poorly informed at the time of purchase even if the shape of the hazard function of car ownership proved to be horizontal. Hence, a perhaps initially rising but then falling hazard function is sufficient but not necessary to establish that car buyers are subject to quality uncertainty at the time of purchase.

We also investigate how *observed* quality differences affect the hazard function. In so

doing, we distinguish between four levels of car quality: new vehicles, used cars submitted by dealers for inspection up to three months before purchase, those privately submitted up to three months prior to sale, and those not inspected. Based in part on our findings from the last section, quality should vary in the same order as the categories are listed, with non-inspected used cars being the worst in the lot. The quality of non-inspected cars can be expected to be the lowest for two reasons. For one, the authorities inform us that most used cars sold uninspected are privately traded. For another, the results in the last section indicate that privately sold cars submitted for inspection have an above-average likelihood of containing a defect. There is no reason to believe that this does not apply equally well to privately sold cars not submitted for inspection. Since those inspected had any discovered defects repaired, they should be of higher quality on average than those not submitted for inspection. The quality of cars sold by dealers can be expected to be even higher on average: for one, that is what the results in the last section show,<sup>16</sup> and for another, dealers have a reputation to lose and thus should make a greater effort to satisfy customers.

Note that buying a new or an inspected used car does not eliminate buyer uncertainty, it merely reduces it. More precisely, it merely increases the expected average quality  $\bar{\mu}$  and lowers  $\varepsilon_t$ . Vehicle inspectors may, for example, not discover all defects. In fact, even a batch of new cars can contain lemons, which is why we consider these vehicles as well. In short, improved observed quality should lower the hazard function, yet not necessarily to zero.

Our investigation of the shape of the hazard function of car ownership is based on the ownership biographies of all 1985 cars registered in the canton of Basle-City between 1985 and 1991. The car histories were constructed from the records of the Basle-City Department of Motor Vehicles. Our observations consist of all spells of ownership these cars generated in Basle-City. An ownership spell begins with the purchase of a car and ends with its being re-sold, where a sale is defined as a situation in which a car with a given chassis number is re-registered to a person with a different social security number. Ownership spells need not end in a sale, however. They could still be in progress at the end of the sample period, or owners could junk their cars or move from the canton before the sample period ends. Such spells are right censored and treated appropriately in the estimation. Spells of ownership begun outside the canton, on the other hand, are left censored and excluded from the sample, as are spells of

---

<sup>16</sup> One could however argue that since any defects discovered are repaired before sale, dealer-submitted cars are on equal footing with privately submitted vehicles.

dealer ownership.<sup>17</sup> In total, our sample consists of 1,161 spells of ownership pertaining to used cars and 2,645 spells applying to new cars. We concentrate on the ownership spells of used cars, employing the ownership spells of new vehicles solely for purposes of comparison.

We apply three types of duration models in our analysis: non-parametric, semi-parametric and parametric approaches. In general, non-parametric models do not impose parametric restrictions on the functional form of the hazard function nor do they restrict the way in which covariates affect the hazard function. Semi-parametric approaches only place parametric strictures on the covariate structure, whereas parametric approaches employ both types of constraints. Parametric models are statistically more efficient, but run the risk of mis-specification, while the opposite applies to non-parametric models. By employing all three approaches, we hope to increase the robustness of our results.

Our non-parametric and semi-parametric approaches are based on grouped durations and a common model setup. The model assumes that the hazard rate, or probability,  $p_i(t)$  that an ownership spell  $i$  that has lasted to at least time  $t-1$  will end in a sale in the interval  $[t-1, t)$  can be sufficiently approximated by the following function<sup>18</sup>

$$p_i(t) = \Phi(\alpha_{0t} + \beta' \mathbf{x}_i + \gamma' \mathbf{z}_{it}) , \quad (3)$$

where  $\Phi(\cdot)$  = standard normal cumulative density function,

$\alpha_{0t}$  = interval-specific constant,

$\mathbf{x}_i$  = vector of constant spell-specific vehicle characteristics,

$\mathbf{z}_{it}$  = vector of time-varying spell-specific vehicle characteristics,

$t$  = 1, ..., T.

We estimate the parameters of the model with the maximum likelihood procedure. To keep the estimation problem manageable, we group the ownership spell durations, which are measured in days, into 26 (= T) equal-length intervals equivalent to one quarter, or 90 days. For convenience, we assume that right censoring occurs at the end of an interval. These assumptions lead to the following likelihood function

$$L = \prod_{t=1}^T \prod_{i=1}^{I_{t-1}} p_i(t)^{d_{it}} [1 - p_i(t)]^{1-d_{it}} , \quad (4)$$

<sup>17</sup> In other words, periods in which a vehicle is registered to a dealer, say, as a demonstration car are not counted as ownership spells and thus excluded from our sample.

<sup>18</sup> KIEFER (1990) discusses alternative grouped duration approaches.

where  $I_{t-1}$  represents the number of spells of car ownership still in progress at the beginning of the duration interval  $[t-1, t)$  and  $d_{it}$  is a dummy variable indicating whether spell  $i$  ended in a re-sale in this interval ( $d_{it} = 1$ ) or not ( $d_{it} = 0$ ). Note that each spell contributes several observations to (4), the number of observations per spell equaling the number of time intervals the spell enters. Applying (4) necessitates splitting each spell into a corresponding number of sub-spells.

Our non-parametric model results from simply dropping the covariates in (3), which is essentially equivalent to regressing the probability of a re-sale on a complete set of interval-specific dummy variables. The maximum likelihood estimates of the coefficients  $\alpha_{0t}$ ,  $t = 1, \dots, T$ , of these dummy variables correspond exactly to non-parametric life-table estimates where right censoring is assumed to occur at the end of an interval.

We choose our grouped-duration approach over the non-parametric Kaplan-Meier and semi-parametric Cox approaches for ungrouped data for computational convenience: it is a simple matter to switch from the non-parametric to the semi-parametric specification, and the nested framework is more conducive to testing. Moreover, the only difference between the Kaplan-Meier estimator and life-table analysis is that the latter imposes a fixed interval length on the spells, whereas the former lets the interval length vary according to the distribution of spell lengths. Furthermore, our approach provides a simple means of including covariates within the same model framework without having to impose the proportional hazard assumption of the Cox model or to correct for ties. As we will see below, our approach does not appear to bias the results.

We measure the explanatory variables in (3) as deviations from their respective means<sup>19</sup> so that in the semi-parametric specification  $\Phi(\alpha_{0t})$  represents the hazard rate of an average spell of ownership in the interval  $[t-1, t)$ , holding the value of all variables constant. Note that the unknown parameter vectors  $\beta$  and  $\gamma$  do not vary across spell intervals. This restriction is necessitated by the fact that many of the cars re-sold in a given interval are unique with regard to one or more of the right-hand variables in (3). In this case, specifying interval-specific coefficients leads to a breakdown of the maximum likelihood estimation procedure.<sup>20</sup>

Our parametric approach is based on non-grouped durations, instead, and rests on the following specification

$$h_i(\tau) = h_0(\tau)\exp(\delta' x_i) \quad , \quad (5)$$

<sup>19</sup> The means apply to all ownership spells, i.e., to those of cars bought new and used taken together.

<sup>20</sup> See GREENE (2000), p. 832.

where  $\tau$  measures continuous time.  $h_0(\cdot)$  is a baseline hazard function of a given parametric form. In contrast to (3), (5) assumes that the effect of covariates is multiplicatively separable from the baseline hazard function (proportional hazard assumption), which in (3) corresponds to  $\Phi(\alpha_{0t})$ ,  $t = 1, \dots, T$ . Accordingly,  $\delta$  does not contain a constant term. The effect of regressors, as specified in (5), is to shift the hazard function upwards or downwards by a factor of proportionality equal to  $\exp(\delta'x_i)$ . Note that we exclude time-varying covariates in the parametric approach.

We assume two different parametric specifications for  $h_0(\tau)$ : a Weibull model, where

$$h_0(\tau) = \lambda p(\lambda\tau)^{p-1} \quad (6)$$

and a log-logistic model, where

$$h_0(\tau) = \lambda p(\lambda\tau)^{p-1}/[1+(\lambda\tau)^p]. \quad (7)$$

$\lambda$  is a location parameter, and  $p$  a scaling parameter. In the Weibull model, the hazard function is monotone increasing for  $0 < p < 1$ , constant for  $p = 1$ , and monotone increasing for  $p > 1$ . In the log-logistic model, on the other hand, the hazard function is monotone decreasing except when  $p > 1$ , in which case it first increases to point  $\tau = (p-1)^{-p}/\lambda$  and decreases monotonically thereafter. Estimates of  $p$  thus provide a means of testing the shape of the hazard function.

In the Weibull model, we also control for unobserved spell heterogeneity. As is well known, unobserved heterogeneity may falsely suggest that a hazard function is decreasing. To control for this bias, we condition the baseline survivor function  $S_0(\tau)$  on a random variable  $v$ . Assuming that  $v$  has a gamma distribution with mean 1 and variance  $\theta$  leads to the following formula for the unconditional baseline hazard function<sup>21</sup>

$$h_0(\tau) = \lambda p(\lambda\tau)^{p-1}[S_0(\tau)]^\theta \quad (8)$$

As is easily seen, (8) reverts to (6) for  $\theta = 0$ , in which case unobserved heterogeneity does not present a problem.

We employ the maximum likelihood procedure as well to estimate the parametric

---

<sup>21</sup> See GREENE (2000), p. 947 for a derivation of (6). This approach is common in parametric modeling, but according to HECKMAN and SINGER (1984) may overparameterize the survivor function leading to serious errors in inference.

model. The likelihood function in this case has the well-known Tobit form, where  $d_i$  indicates whether a spell of car ownership ended in a sale ( $d_i = 1$ ) or is right censored ( $d_i = 0$ ).

$$\begin{aligned}
 L &= \prod_{\tau=1}^T f(\tau_i)^{d_i} S(\tau_i)^{1-d_i} \\
 &= \prod_{\tau=1}^T h(\tau_i)^{d_i} S(\tau_i)
 \end{aligned}
 \tag{9}$$

In contrast to (4), each spell contributes just one observation to the likelihood function in (9).

*Table 3* reports the means of the ownership spells in our sample, distinguishing between uncensored ("re-sold") and right censored ("not re-sold") spells. Ownership duration measures the length of time (in quarters) that a vehicle had been held on average at the time the ownership spell ended, either through a sale or censoring. The remaining variables constitute the right-hand variables ( $\mathbf{x}$ ,  $\mathbf{z}$ ) in our model. With the exception of "age at purchase", all right-hand variables represent dummy variables. In the semi-parametric model, we replace "age at purchase" with the current age of the car in each time interval, which constitutes a time-varying variable, the sole regressor of this kind in the study.

The inspection variables indicate whether a car was submitted to a safety inspection in the last three months prior to purchase, i.e., before the new ownership spell began, and in addition, whether the car was brought for inspection by a dealer ("dealer"). As mentioned before, the inspection variable serves as a measure of observed vehicle quality. The higher the quality of a vehicle is on average, the less likely that buyers should be negatively surprised and thus attempt to re-sell their cars in the near future. Indeed this is what *Table 3* suggests as a smaller share of the used cars re-sold had been inspected up to three months prior to purchase: 28.8 percent compared to 32.9 percent for cars not re-sold. The relative difference is even more pronounced among those cars that dealers submitted for inspection: 18.8 versus 22.4 percent.

The dummy variable "company owned" indicates whether a car was registered to a firm rather than to a person during a particular spell of ownership. As can be seen, no firm purchased a used car in our sample. "Commercially used" denotes that the automobile was being used as a rental car, a taxi, or for driver education. The brands listed in the table are those for which dummy variables were created and appear in their or-

der of prominence in our sample. The reference group is made up of all other makes.

*Table 3: Sample Means of Ownership Spells*

Variable	Used Cars		New Cars	
	Re-Sold	Not Re-Sold	Re-Sold	Not Re-Sold
Total Number of Spells	160	1001	465	2180
Ownership Duration (in quarters)	3.644	5.907	16.98	22.48
Age at Purchase (in quarters)	18.05	19.99	0.000	0.000
Inspected up to 3 Mo. before Purchase	0.288	0.329	--	--
Inspected up to 3 Mo. before Purchase (dealer)	0.188	0.224	--	--
Company Owned	0.000	0.000	0.002	0.002
Commercially Used	0.013	0.015	0.009	0.010
Diesel Engine	0.013	0.030	0.024	0.036
44 Brands in total:				
Opel	0.088	0.113	0.101	0.154
VW	0.063	0.092	0.086	0.128
Toyota	0.138	0.084	0.112	0.094
Renault	0.075	0.073	0.065	0.047
Ford	0.038	0.061	0.071	0.050
Fiat	0.069	0.058	0.067	0.039
Honda	0.044	0.040	0.043	0.053
Audi	0.038	0.043	0.054	0.048
Citroen	0.019	0.044	0.052	0.040
BMW	0.019	0.046	0.039	0.042
Peugeot	0.031	0.040	0.034	0.042
Alfa	0.069	0.056	0.047	0.026
Nissan	0.063	0.029	0.041	0.038
Mitsubishi	0.063	0.036	0.043	0.033
Mazda	0.025	0.029	0.026	0.035
Mercedes	0.031	0.036	0.019	0.032
Volvo	0.013	0.026	0.022	0.020
Suzuki	0.050	0.021	0.013	0.013
Subaru	0.031	0.011	0.006	0.018
Lancia	0.013	0.020	0.026	0.011
Saab	0.000	0.005	0.006	0.007
Porsche	0.006	0.005	0.004	0.005
Seat	0.000	0.003	0.004	0.005
Other	0.019	0.030	0.019	0.019

*Table 3* suggests that re-sold cars ("uncensored") are held for shorter spells on average and tend to be somewhat newer, not in commercial use, and equipped with a gasoline engine. The table further indicates that automobiles from Alfa, Fiat, Mitsubishi, Nissan, Renault and Toyota are generally more likely to be traded, whether bought new or used, while German cars (Audi, BMW, Ford, Mercedes, Opel, Porsche, VW), which have a reputation of being particularly solid, are often more likely to be held

back from the market.

*Table 4* presents our regression results. The first three columns present the results yielded by the non-parametric model. The first column considers all used cars together, the second and third contain separate estimates for used cars not submitted up to three months prior to purchase and for those that were, respectively. The estimates of the semi-parametric model, which includes covariates, follow in the next set of columns. Four specifications were estimated: one applying just to new cars. The used car estimates differ with respect to the variables controlling for a prior inspection. The first of the semi-parametric specifications contains a dummy variable controlling for whether a car had been submitted to a safety inspection up to three months before purchase, the second interacts this dummy with another dummy controlling for whether a dealer submitted the vehicle for inspection ("dealer-initiated inspection"), and the third specification includes both variables. The final columns pertain to the estimates of the parametric models, which unlike the other models are not based on grouped durations.

The estimates of the interval-specific constants  $\alpha_{0t}$  ( $t = 1, \dots, T$ ) contained in the non-parametric and semi-parametric model (3) and which mark the ownership quarter (Q) in which a car was re-sold, if at all,<sup>22</sup> yield consistent estimates  $\Phi(\hat{\alpha}_{0t})$  of the probability  $p(t)$  that an owner of a used car will sell his or her vehicle in interval  $[t-1, t)$ , given that he or she has not done so by time  $t-1$ . These estimated hazard rates provide non-parametric and semi-parametric estimates of the underlying hazard function, respectively. The semi-parametric estimates of the hazard function, unlike the non-parametric estimates, hold the values of all covariates constant at their mean values across all spells, both those of used and new cars. Hence, the hazard rates apply to a car with constant average observable attributes. The parametric models, on the other hand, identify the hazard function by first assuming it exhibits a specific parametric form (Weibull, log-logistic) and then estimating the parameters  $\lambda$  and  $p$ .

The hazard function estimates in *Table 4* address two issues:

- Does the hazard function of car ownership, after perhaps initially increasing, decline with the duration of ownership as the quality uncertainty on the part of buyers would lead one to expect?
- What effect, if any, does higher average vehicle quality have on the hazard func-

---

<sup>22</sup> Missing constant estimates indicate that no car in our sample was sold in the given quarter of ownership. In these cases it is impossible to estimate the associated hazard rate because  $\alpha_{0t}$  equals  $-\infty$ .

tion?

We turn first to the first issue.

Our regression results in general support the prediction of a declining hazard function. To start, the likelihood ratio statistic of 19.79 pertaining to the non-parametric estimates appearing in the first column of *Table 4* rejects the null hypothesis of a uniform hazard rate across all spell intervals at the 7.1 percent level of significance, implying that the interval-specific hazard rates do indeed differ.<sup>23</sup> Moreover, an eyeball inspection of the hazard rates which the estimates of the interval-specific constants of this specification imply (*Figure 1*) suggests that the hazard function is not only not flat, but roughly decreasing as well. This is further supported by *Figure 2*, which presents a graph of the integrated hazard function across the first 400 days<sup>24</sup> of car ownership. The integrated hazard, as the name implies, simply sums the hazard rates cumulatively across the individual spells lengths ranked in ascending order. If the hazard rates were uniform, this summing would yield a straight line. The fact that the curve is concave from below, instead, indicates that the hazard rates are decreasing as *Figure 1* suggests. Moreover the log of the integrated hazard, when plotted against the log of ownership duration (*Figure 3*), yields a roughly straight line with a slope of 0.76, suggesting that a Weibull model with a parameter  $p$  equal to 0.76 would provide a good approximation of the underlying hazard function.<sup>25</sup> Our estimates of the parametric models support that conjecture. As the last columns in *Table 4* indicate, the Weibull and the log-logistic models yield estimates for  $p$  of 0.800 and 0.843, respectively, which are significantly different from one at the one percent level (one-sided test), implying that the baseline hazard function is monotonically decreasing throughout. Note that in contrast to the non-parametric model, the parametric specifications do not ignore spell heterogeneity,<sup>26</sup> nor do they group the spell durations into three-month interval classes, both factors which could bias the non-parametric results. Both the parametric and the non-parametric results suggest that the probability of car being re-sold after purchase declines by around 37 percent within the first year of ownership.

The semi-parametric results, on the other hand, are less clear on the form of the hazard function. The estimates pertaining solely to cars bought new support the hypothe-

---

<sup>23</sup> The value of the restricted log-likelihood function  $\ln L(\beta_0)$  in the non-parametric and semi-parametric models refers throughout to a specification in which the interval-specific constants are replaced by a single constant.

<sup>24</sup> Note that the integrated hazard function is based on non-grouped durations.

<sup>25</sup> The survivor function of the Weibull model presented in (6) has the form  $S(\tau) = \exp(-\lambda\tau)^p$ . Taking double logs of both sides yields  $\ln[-\ln S(\tau)] = \ln[\Lambda(\tau)] = p \cdot \ln \lambda + p \cdot \ln \tau$ , where  $\Lambda(t)$  corresponds to the integrated hazard.

<sup>26</sup> Note that the estimate for  $\theta$  suggests that unobserved heterogeneity is not a problem.

sis of differing interval-specific re-sale probabilities. The likelihood ratio statistic of 499.22 rejects the null hypothesis of a uniform hazard rate across all duration intervals at the one percent level. The hazard rates that the estimated interval-specific constants imply appear in *Figure 4*, where they are compared with those for used cars.<sup>27</sup> As can be seen, the hazard rates of cars bought new do not suggest any simple parametric form, which is why we refrain from estimating them parametrically. After increasing in the first two quarters of ownership, the conditional re-sale probabilities for new cars decline over the next six quarters before rising again after about two years.<sup>28</sup> This later increase undoubtedly stems largely from changes in owners' tastes or the constraints facing them. It is interesting to note that even new cars face a non-zero probability of being re-sold within the first six months of ownership, implying that buyers of new cars are also subject to quality uncertainty. The degree of quality uncertainty is undoubtedly greater for buyers of used cars, however, which would explain why their conditional re-sale probabilities are much higher. Note too that the re-sale probability pattern for new and used cars is roughly similar from the 9th to 14th quarter of ownership, suggesting that owners are equally likely to re-sell new or used cars once differences in quality uncertainty have been eliminated.

The other semi-parametric regression results, pertaining solely to used cars, and the remaining non-parametric findings, applying separately to inspected and non-inspected used cars, are unable to reject the null hypothesis of a uniform hazard rate across all duration intervals at commonly acceptable levels of significance, although the estimates of the interval-specific constants nevertheless all point to a declining hazard function similar to the one in *Figure 1*. Apparently our sample is not sufficiently rich to yield unambiguous estimates of the hazard function of used cars when simultaneously controlling for the observed heterogeneity of spells.

The results in *Figure 4* bring us to the second issue: the effects of observed car quality on the hazard function of car ownership. As new cars are of higher quality on average, the results applying to new and used cars already suggest that improved quality shifts

---

<sup>27</sup> The coefficient estimates upon which the hazard rates in *Figure 4* are based do not appear in *Table 4*. They result from an estimation of the semi-parametric model for used-car spells in which the inspection variables have been dropped. Hence, the hazard rates for new and used cars presented in *Figure 4* both control for the same covariate effects. As all regressors, save the inspection variables, are measured in deviations from their respective overall (used and new combined) sample means, both sets of hazard rates pertain to the same constant average profile of characteristics and thus can be directly compared.

<sup>28</sup> The size of the estimated standard errors do not allow one to reject the hypothesis of uniform hazard rates across the first four quarters of car ownership, however, although they definitely fall thereafter.

the hazard function downwards.<sup>29</sup> We find this to hold generally for used cars as well. Most results pertaining to used vehicles indicate that a prior inspection lowers the hazard function, implying that previously inspected cars are less likely to be re-sold. That is particularly true if a dealer submitted the car for inspection ("dealer-initiated inspection"). As we saw in the previous section, used cars presented for inspection by a dealer are less likely to have a defect.

The non-parametric estimates are somewhat less clear on the impact of higher observed mean quality. The estimated hazard rates appearing in *Figure 5*, which stem from separate regressions for inspected and non-inspected used cars,<sup>30</sup> suggest that a prior inspection generally lowers the conditional probability of a re-sale. Yet the results from nonparametric homogeneity tests, which - roughly speaking - compare the observed number of sales in each ownership interval with the corresponding expected number of sales were both hazard functions identical, are inconclusive.<sup>31</sup> The so-called log-rank test rejects the null hypothesis of identical hazard functions at a 7.7 percent level of significance, while the generalized Wilcoxon test can only reject it at a generally unacceptably high level of 31.3 percent. Since the log-rank statistic is more sensitive to differences at high durations, and the generalized Wilcoxon statistic to those at low durations, our test results imply that the hazard functions of inspected and non-inspected used cars differ more at higher durations, which *Figure 5* also suggests.

The semi-parametric and parametric results, on the other hand, are less ambiguous, in part, of course, because they impose additional structure on the way covariates affect the hazard function. The parametric models employed here assume a strictly proportional effect across all ownership durations, whereas the semi-parametric specification allows a bit more flexibility. Despite these differences, both approaches indicate that a prior inspection lowers the probability of a re-sale by almost 30 percent on average. The associated hazard rates appear in *Figures 6 to 8*. The decrease is even greater if only "dealer-initiated" inspections are considered. In this case the hazard function falls by around 35 percent on average. Unfortunately, a separate effect for dealer and non-dealer-initiated inspections could not be established at commonly acceptable levels of statistical significance. We nevertheless present the results of including both variables within the semi-parametric framework in *Figure 7* to convey an impression of how

---

<sup>29</sup> Note that by running separate regressions for new and used cars we control for any possible interaction or slope effects.

<sup>30</sup> These appear in the second and third column of estimates in *Table 4*. By running separate regressions we again control for interaction effects.

<sup>31</sup> KALBFLEISCH and PRENTICE (1980) describe the tests in more detail.

strong the two effects could be.

In summary, our results suggest that the hazard function of car ownership is indeed decreasing, as quality uncertainty would lead one to expect, and that higher average observable quality lowers the probability of a re-sale, yet without eliminating it.

On the basis of these results, it seems reasonable to assert that informational asymmetry between sellers and buyers prevails in cases in which used cars were sold without being submitted to a safety inspection within three months of being sold. This applies to roughly two-thirds of the used cars in our sample (*Table 3*). Uninspected cars are usually sold privately, and according to the results pertaining to adverse selection, privately sold cars submitted to a safety inspection are more likely to have a defect. It seems fair to reason that this probably holds even more so for non-inspected cars.

It might even be reasonable to maintain that private sellers of used cars that were inspected beforehand enjoy an informational advantage over buyers, as the disproportionate number of their cars that fail to pass inspection implies that sellers of such cars know something is amiss with their vehicles, especially given that similar cars that dealers present for inspection are less likely to fail. This reasoning implies that owners are also aware of defects that the inspectors failed to discover.

Although present, the lemons problem does not appear to be widespread, however. For one, 254 or almost 22 percent of the 1,161 used cars in our sample were submitted for inspection by dealers before sale and thus presumably purchased from them. In these cases, informational asymmetry can probably be ruled out. For another, only 160 or roughly 10 percent of the used cars in our sample were re-sold after purchase, and of these 110 or roughly 70 percent were re-sold within one year of purchase. On the other hand, there may have been other dissatisfied buyers for whom the transaction costs of re-selling their cars were too high. Hence, informational asymmetry may be more prevalent than the frequency of re-sales indicate.

Table 4: Hazard Function Estimates

Explanatory Variable	Non-Parametric Models			Semi-Parametric Models				Parametric Models	
	Used Cars all	Used Cars non-inspected	Used Cars inspected	Used Cars all	Used Cars all	Used Cars all	New Cars all	Used Cars Weibull	Used Cars Log-Logistic
Q1	-1.808*** (0.070)	-1.788*** (0.083)	-1.852*** (0.127)	-1.760*** (0.147)	-1.766*** (0.146)	-1.755*** (0.147)	-3.391*** (0.278)	--	--
Q2	-1.873*** (0.080)	-1.851*** (0.096)	-1.920*** (0.143)	-1.826*** (0.157)	-1.834*** (0.155)	-1.822*** (0.157)	-2.919*** (0.143)	--	--
Q3	-1.912*** (0.090)	-1.860*** (0.106)	-2.030*** (0.168)	-1.863*** (0.163)	-1.870*** (0.162)	-1.859*** (0.163)	-3.384*** (0.277)	--	--
Q4	-1.991*** (0.104)	-2.046*** (0.137)	-1.904*** (0.163)	-1.947*** (0.176)	-1.952*** (0.175)	-1.941*** (0.176)	-3.386*** (0.278)	--	--
Q5	-2.013*** (0.115)	-1.980*** (0.140)	-2.080*** (0.203)	-1.946*** (0.187)	-1.954*** (0.186)	-1.942*** (0.187)	-3.200*** (0.210)	--	--
Q6	-2.013*** (0.125)	-2.011*** (0.157)	-2.017*** (0.207)	-1.945*** (0.195)	-1.953*** (0.194)	-1.941*** (0.195)	--	--	--
Q7	-1.898*** (0.125)	-1.715*** (0.139)	-2.500*** (0.353)	-1.823*** (0.199)	-1.829*** (0.198)	-1.818*** (0.200)	--	--	--
Q8	-2.261*** (0.191)	-2.066*** (0.204)	--	-2.168*** (0.247)	-2.170*** (0.246)	-2.160*** (0.247)	--	--	--
Q9	-2.024*** (0.169)	-1.978*** (0.210)	-2.100*** (0.285)	-1.943*** (0.234)	-1.947*** (0.233)	-1.936*** (0.234)	-3.064*** (0.174)	--	--
Q10	-2.610*** (0.341)	-2.406*** (0.364)	--	-2.557*** (0.390)	-2.561*** (0.389)	-2.550*** (0.390)	-2.908*** (0.142)	--	--
Q11	-2.513*** (0.352)	--	-2.222*** (0.386)	-2.396*** (0.386)	-2.413*** (0.388)	-2.395*** (0.388)	-2.329*** (0.074)	--	--
Q12	--	--	--	--	--	--	-2.325*** (0.074)	--	--
Q13	-2.246*** (0.383)	-1.971*** (0.421)	--	-2.187*** (0.434)	-2.191*** (0.433)	-2.179*** (0.434)	-2.164*** (0.065)	--	--
Q14	-2.100*** (0.402)	-1.819*** (0.444)	--	-1.998*** (0.441)	-2.001*** (0.440)	-1.989*** (0.441)	-2.133*** (0.065)	--	--
Q15	--	--	--	--	--	--	-2.111*** (0.065)	--	--
Q16	--	--	--	--	--	--	-2.056*** (0.064)	--	--
Q17	--	--	--	--	--	--	-2.137*** (0.070)	--	--

Table 4: Hazard Function Estimates (cont'd.)

Explanatory Variable	Non-Parametric Models			Semi-Parametric Models				Parametric Models		
	Used Cars all	Used Cars non-inspected	Used Cars inspected	Used Cars all	Used Cars all	Used Cars all	New Cars all	Used Cars Weibull	Used Cars Log-Logistic	
Q18	--	--	--	--	--	--	-2.271*** (0.082)	--	--	
Q19	--	--	--	--	--	--	-2.070*** (0.069)	--	--	
Q20	--	--	--	--	--	--	-2.027*** (0.068)	--	--	
Q21	--	--	--	--	--	--	-2.102*** (0.075)	--	--	
Q22	--	--	--	--	--	--	-2.344*** (0.096)	--	--	
Q23	--	--	--	--	--	--	-2.178*** (0.084)	--	--	
Q24	--	--	--	--	--	--	-2.220*** (0.089)	$\lambda \times 10^3$	0.155*** (0.052)	0.207*** (0.016)
Q25	--	--	--	--	--	--	-2.186*** (0.090)	$\rho$	0.800*** (0.090)	0.843*** (0.063)
Q26	--	--	--	--	--	--	-2.490*** (0.145)	$\theta$	0.083 (1.549)	--
Inspected Prior to Purchase	--	--	--	-0.143* (0.077)	--	-0.062 (0.117)	--	-0.432* (0.245)	-0.419* (0.250)	
Dealer-Initiated Inspection	--	--	--	--	-0.173** (0.088)	-0.120 (0.134)	--	--	--	
Age	--	--	--	-0.003 (0.009)	-0.004 (0.009)	-0.004 (0.009)	--	0.021 (0.028)	0.015 (0.029)	
Company Owned	--	--	--	--	--	--	0.105 (0.412)	--	--	
Commercially Used	--	--	--	-0.013 (0.303)	-0.022 (0.305)	-0.020 (0.304)	0.125 (0.191)	-0.202 (1.065)	-0.241 (1.152)	
Diesel	--	--	--	-0.201 (0.289)	-0.212 (0.291)	-0.206 (0.290)	-0.196* (0.117)	-0.652 (1.073)	-0.696 (1.111)	
Observations	6,301	3,938	1,977	6,301	6,301	6,301	56,136	1,161	1,161	
$\ln L(\beta_0)$	-745.68	-516.15	-218.45	-725.98	-725.71	-725.50	-2667.38	-645.06	-644.54	
$\ln L(\beta^*)$	-735.78	-511.43	-215.77	-718.14	-717.89	-717.74	-2417.77	-626.91	-627.10	
$-2[\ln L(\beta_0) - \ln L(\beta^*)]$	19.79*	9.43	5.37	15.68	15.65	15.51	499.22***	36.29*	34.87*	
$LRI [1 - \ln L(\beta^*) / \ln L(\beta_0)]$	0.013	0.009	0.012	0.011	0.011	0.011	0.094	0.028	0.027	

Asterisks denote statistical significance at the 10% (\*), 5% (\*\*), or 1% (\*\*\*) level. Estimated standard errors appear in parentheses.

Figure 1: Hazard Function of Used-Car Ownership: Non-Parametric Model

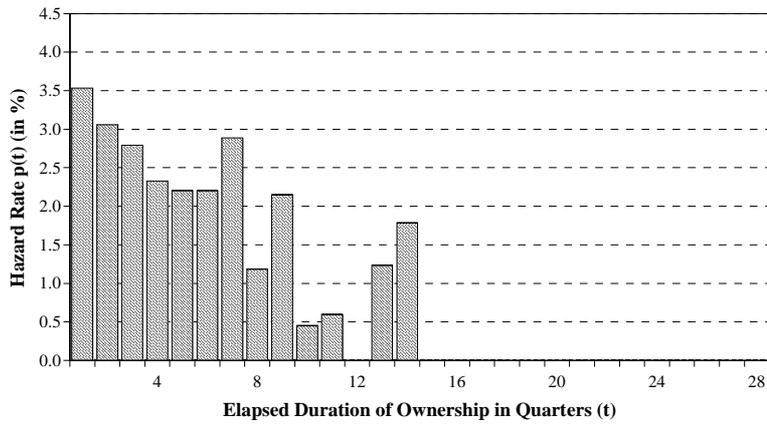


Figure 2: Integrated Hazard Function of Used-Car Ownership: Non-Parametric Model

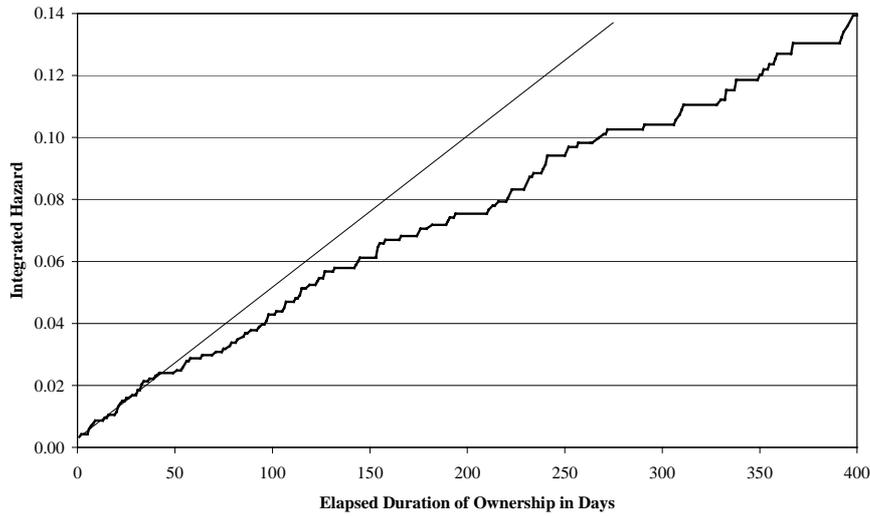


Figure 3: Log Integrated Hazard Function of Used-Car Ownership: Non-Parametric Model

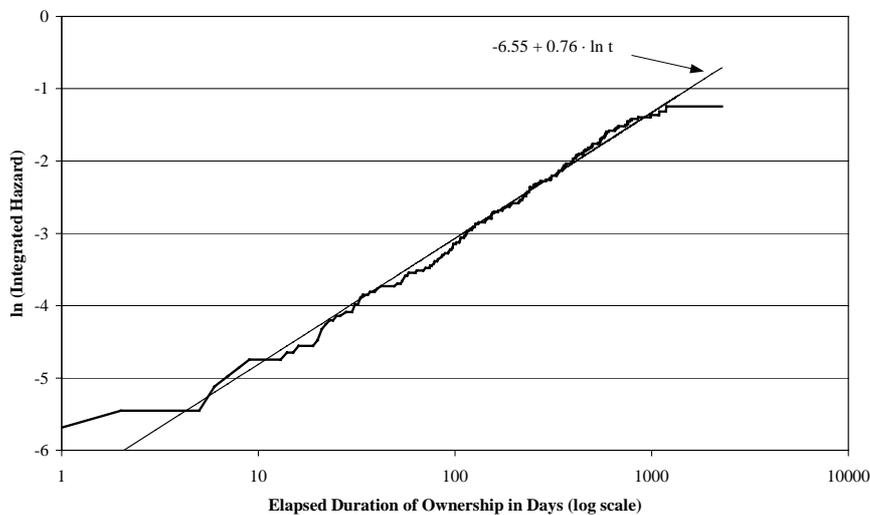


Figure 4: Hazard Function of New and Used-Car Ownership: Semi-Parametric Model

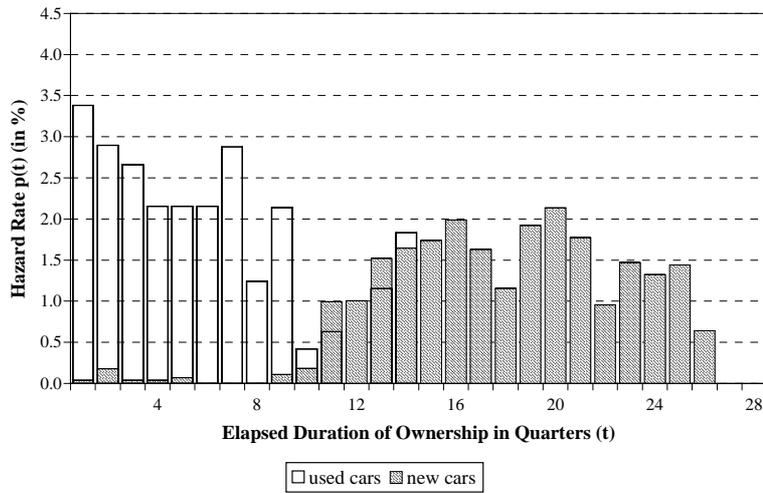


Figure 5: Hazard Function of Used-Car Ownership: Non-Parametric Model

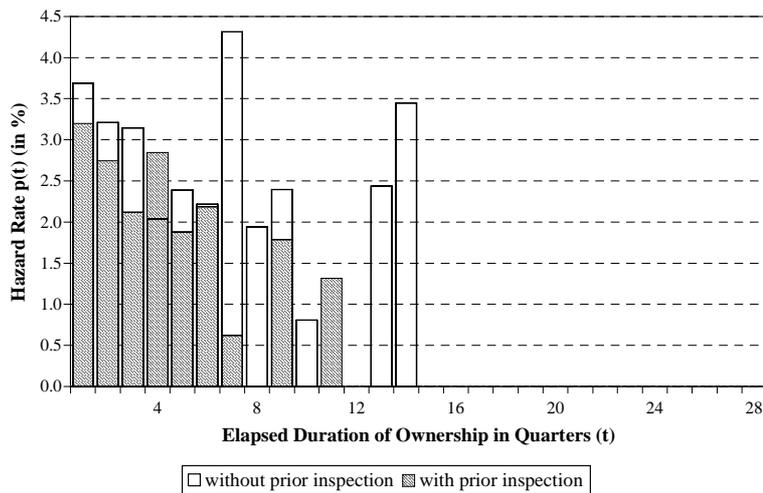


Figure 6: Hazard Function of Used-Car Ownership: Semi-Parametric Model

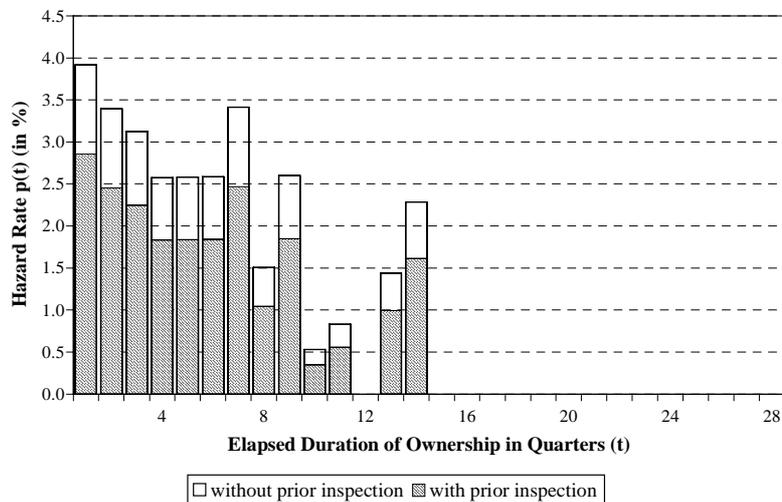


Figure 7: Hazard Function of Used-Car Ownership: Semi-Parametric Model

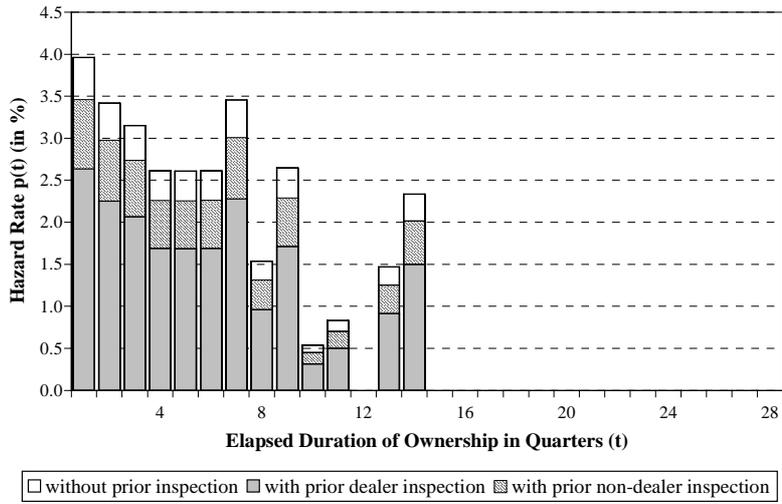
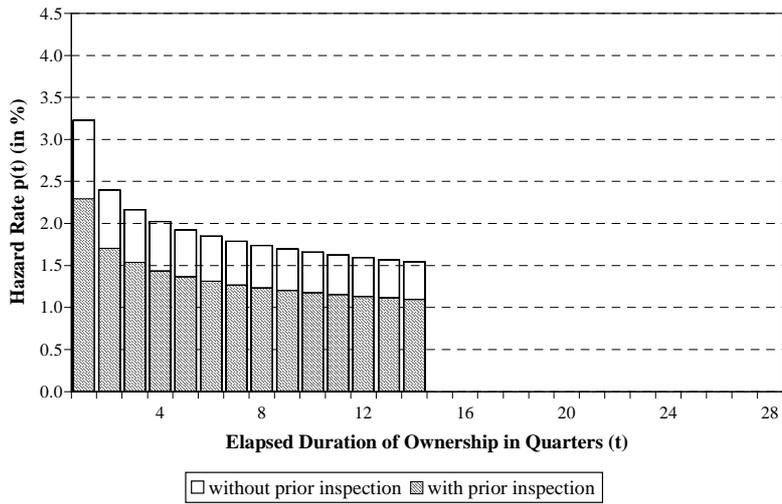


Figure 8: Hazard Function of Used-Car Ownership: Parametric Model (Weibull)



## 6. Conclusions

This study has attempted to extend the work of previous empirical studies of the lemons model by investigating the presence of quality uncertainty among buyers along with testing for adverse selection among goods sellers choose to trade. We applied our approach to a sample consisting of all 1985-make cars that were registered in the Swiss canton of Basle-City sometime in the period 1985-1991.

Our results pertaining to adverse selection indicated that cars sold privately are indeed more likely to have a defect than a randomly chosen vehicle, implying that owners of poorer quality cars knowingly put them up for sale. At the same time, our estimates relating to the quality uncertainty afflicting buyers showed that these individuals are not fully aware of what they are purchasing, as evidenced by our finding that the probability that buyers will re-sell their vehicles is higher in the first months of ownership than latter. This holds true whether a car was bought new or used, or whether it was inspected before purchase or not. However, the higher the mean quality of a car is, the less likely it will be quickly re-sold, supporting our conjecture that quality uncertainty lies behind buyers' propensity to re-sell their vehicles shortly after purchase.

Based on our results, we conclude that at least the buyers of used cars not previously inspected were at an informational disadvantage vis à vis sellers. This applied to roughly two-thirds of the used cars purchased in our sample. Less than 10 percent of the used cars purchased were re-sold within the first year of ownership, however, so the lemons problem does not appear to be widespread. On the other hand, more buyers may have been unpleasantly surprised by their purchases, yet shunned the transaction costs that the re-sale of their cars would have entailed. Hence, our study may underestimate the full scope of the lemons problem in the sample.

## References

- AKERLOF, G.: The Market for “Lemons”: Quality Uncertainty and the Market Mechanism, *Quarterly Journal of Economics*, 84 (1970), 488-500
- BOND, E. W.: A Direct Test of the “Lemons” Model: The Market for Used Pickup Trucks, *American Economic Review*, 72 (1982), 836-840
- BOND, E. W.: Test of the “Lemons” Model: Reply, *American Economic Review*, 74 (1984), 801-804
- BORLAND, J., J. LYE: Matching and Mobility in the Market for Australian Football Coaches, *Industrial and Labor Relations Review*, 50 (1996), 143-158
- BULOW, J.: Durable Goods Monopolists, *Journal of Political Economy*, 90 (1982), 314-332
- CHAPMAN, K., L. SOUTHWICK: Testing the Matching Hypothesis: The Case of Major-League Baseball, *American Economic Review*, 81 (1991), 1352-1360
- CHEZUM, B., B. WIMMER: Roses or Lemons: Adverse Selection in the Market for Thoroughbred Yearlings, *Review of Economics and Statistics*, 79 (1997), 521-526
- CHIAPPORI, P.-A., B. SALANIÉ: Testing for Asymmetric Information in Insurance Markets, *Journal of Political Economy*, 108 (2000), 56-78
- COHEN, A.: Asymmetric Information and Learning in the Automobile Insurance Market, mimeo Harvard University (2001), [http://papers.ssrn.com/paper.taf?abstract\\_id=28996656-78](http://papers.ssrn.com/paper.taf?abstract_id=28996656-78)
- CUTLER, D., R. ZECKHAUSER: The Anatomy of Health Insurance, Forthcoming in A. Cuyler, J. Newhouse (eds.), *Handbook of Health Economics*, Elsevier Science (2000)
- DIONNE, G., C. GOUÏÉROUX, C. VANASSE: Testing for Evidence of Adverse Selection in the Automobile Insurance Market: A Comment, *Journal of Political Economy*, 109 (2001), 444-453
- EMONS, W., J. VON HAGEN: Asset Prices and Public Information: An Empirical Investigation in the Market for Automobiles, *European Economic Review*, 35 (1991), 1529-1542
- FABEL, O., E. LEHMANN: Adverse Selection and the Economic Limits of Market Substitution: An Application to E-Commerce and Traditional Trade in Used Cars, University of Constance discussion paper no. 302 (2000), [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=213088](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=213088)
- HAUTSCH, N., E. LEHMANN, S. WARNING, B. FRICK: Shirking or Mismatch? Coach-Team Separation in German Professional Soccer, University of Constance discussion paper no. 313 (2001)
- HENDEL, I., A. LIZZERI: Adverse Selection in Durable Goods Markets, *American Economic Review*, 89 (1999), 1097-1115
- HECKMAN, J., B. SINGER, A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data, *Econometrica*, 52 (1984), 271-320
- GENESOVE, D.: Adverse Selection in the Wholesale Used Car Market, *Journal of Po-*

- litical Economy*, 101 (1993), 644-665
- GIBBONS, R., L. KATZ: Layoffs and Lemons, *Journal of Labor Economics*, 9 (1991), 351-380
- GREENWALD B., R. GLASSPIEGEL: Adverse Selection in the Market for Slaves: New Orleans 1830-1860, *Quarterly Journal of Economics*, 98 (1983), 479-499
- GREENE, W., *Econometric Analysis*, 4th ed., Prentice Hall: London (2000)
- JOVANOVIC, B.: Job Matching and the Theory of Turnover, *Journal of Political Economy*, 87 (1979), 972-990
- KALBFLEISCH, J., R. PRENTICE, *The Statistical Analysis of Failure Time Data*, John Wiley & Sons: New York (1980).
- KIEFER, N.: Economic Duration Data and Hazard Functions, *Journal of Economic Literature*, 26 (1988), 646-679
- KIEFER, N.: Econometric Methods for Grouped Duration Data, J. Hartog, G. Ridder, J. Theeuwes (eds.), *Panel Data and Labor Market Studies*, North-Holland: Amsterdam (1990), 97-117
- LANCASTER, T., *The Econometric Analysis of Transition Data*, Cambridge University Press: Cambridge (1990)
- LANCASTER, T., G. IMBENS, P. DOLTON: Job Separations and Job Matching, R. Heijmans, H. Neudecker (eds.), *The Practice of Econometrics. Studies on Demand, Forecasting, Money and Income*, Kluwer Academic Publishers: Dordrecht (1987), 31-43
- LEHN, K.: Information Asymmetries in Baseball's Free Agent Market, *Economic Inquiry*, 22 (1984), 37-44
- OHTAKE, F., Y. OHKUSA: Testing the Matching Hypothesis: The Case of Professional Baseball in Japan with Comparison to the US, *Journal of the Japanese and International Economies*, 8 (1994), 204-219
- PORTER, R., P. SATTLER: Patterns of Trade in the Market for Used Durables: Theory and Evidence, NBER Working Paper 7149 (1999)
- PRATT, M., G. HOFER: Test of the "Lemons" Model: Comment, *American Economic Review*, 74 (1982), 798-800
- PUELZ, R., A. SNOW: Evidence on Adverse Selection: Equilibrium Signaling and Cross Subsidization in the Insurance Market, *Journal of Political Economy*, 102 (1994), 236-257
- ROSENMAN, R., W. WILSON: Quality Differentials and Prices: Are Cherries Lemons?, *Journal of Industrial Economics*, 39 (1991), 649-658
- SHELDON, G.: Dauerhaftigkeit der Beschäftigung in der Schweiz, W. Franz (ed.), *Mikro- und makroökonomische Aspekte der Arbeitslosigkeit*, BeitrAB 165 (1992), 36-73