Durable Goods Lease Contracts and Used-Goods Market Behavior: An Experimental Study

Kay-Yut Chen
HP Labs

Suzhou Huang
Research and Advanced Engineering
Ford Motor Company

March 16, 2004

Abstract

Leasing has become an increasingly prominent way for consumers to acquire durable goods such as automobiles. How markets respond to changes in lease contracts has enormous implications to producers such as Ford Motor Company. In this paper, an experimental model was developed to study the interaction between lease contracts that embed an option to purchase and an underlying used-goods market. Experiments with subjects playing roles of heterogeneous consumers have confirmed many salient features predicted by the theoretical model. These features include the segmentation of subjects into classes of behavior, and directional response to pricing in the used-good market to the provision in lease contracts.

1 Introduction

Leasing has become an increasingly prominent way for consumers to acquire durable goods. Very often, lease contracts embed options that allow lessees the right but not the obligation to purchase the item at the end of the lease. This form of lease contract is very popular in the automobile industry. In this paper, an experimental model was developed to study how this kind of lease contracts interacts with an underlying used-goods market. This research, although self-contained, is the first stage of collaboration between HP Labs and the Ford Motor company to create a general framework to address some of the unique issues in automobile marketing.

The standard option pricing theory approach assumes perfect competition and frictionless market (Black and Scholes 1973, Merton 1973). In this framework, agents are assumed to be homogenous and non-strategic, and transaction costs are all negligibly small. Furthermore, producers are assumed to have little market power and hence are treated as price takers. These assumptions are quite reasonable for financial and well-traded commodity markets. However, they often fail to capture the key features of those markets in which durable goods are leased such as automobiles and heavy machineries. In these durable-goods markets, consumers are heterogeneous and need to act strategically in their consumption decisions in a time-consistent manner due to sizable transaction costs that they have to incur for trading used goods. Similarly, many of these durable goods are often made by a few big producers with a high degree of differentiation. This, in turn, implies that, rather than simply acting as price takers, the producers can enjoy certain market power in pricing their goods and changing provisions in lease contracts. Aiming to gain insights, Huang and Yang (2002) had constructed a theoretical model that explicitly incorporates many of the salient features of the economic environment that is more appropriate for these durable-goods markets.

However, some issues remain unresolved if the purpose is to adapt the insights to make policy decisions in the real world. One key issue is whether the model is robust with respect to the stringent rationality requirements imposed upon the consumers. Huang and Yang (2002) employed the solution concept of the Markov perfect equilibrium (Maskin and Tirole 1988). The solution requires the consumer to have perfect knowledge of the present and future prices, as well as the supply and demand of used goods, which are all endogenously determined by solving complicated mathematical equations. This is obviously beyond the undertaking of an average consumer. Even if every agent in the system can perform the mathematics required, there is ample evidence to show that people are neither risk-neutral nor even adhering to expected utility maximization (Camerer 1995). Another question is whether the theoretical results are robust with respect to variations of the price discovery process, which can vary depending on the particular market mechanisms used. It is also impractical trying to infer the answer from real world data because there are many unobserved or uncontrollable variables. The most promising approach is laboratory experiment.

A series of experiments was conducted at HP Experimental Economics Lab. In each experiment, around 23-28 subjects were recruited to play the role of consumers in a hypothetic durable-goods market. Standard experimental economics procedures were followed while there was a slight variation to the standard design of treatments. We gave subjects exact information about the experiment. They were told that their monetary rewards depended on their aggregate performance of the experiment. We preserved anonymity with respect to roles and payment and we used no deception. The experimental model was directly adapted from the setting in Huang and Yang (2002). There was one brand of homogenous goods when they are new. Each unit of goods was associated with a quality measure. A new unit always started with the quality of one. In the first period, a random amount was consumed. The residual or leftover quality was observed at the end of first period, which later would be consumed in the second period. We chose this particular structure to capture the characteristics of automobile market. Given a brand, new cars are generally identical. Thus, all new units started with the same quality that is normalized to one. The primary measurement to determine the relative value of a used car of the same brand is its usage such as mileage driven. In any given period of time, the actual mileage accrued is uncertain at the inset of a lease contract to the lessee. This is why we chose that the quality consumed in the first period is uncertain ex ante. On the other hand, when a lessee is deciding whether to exercise the option at the lease end, the residual quality is ex post and hence is treated observable to the lessee. For simplicity, we assume that each good has a lifetime of two periods. This implies that any residual quality is completely consumed in the remaining life span of the used good, i.e., the next period after its lease.

Each subject is only allowed to have at most one unit of the good, new or used. A positive value is given to the subject at each period if he or she owns a unit. This value is a function of the quality consumed and a private parameter called the willingness-to-pay. Each subject had a different willingness-to-pay parameter, which was chosen to span uniformly over an interval. To focus on issues related to lease we limit the producer only lease its products, and hence outright selling is left out of the scope of this study. In each period, a subject had four alternatives: start a new lease, purchase a unit from the used market, exercise the option to purchase the leased unit at the end of the term if the subject was a lessee in the preceding period, or hold no unit. If a subject started a new lease, he would consume a random amount (the mean and variance of this amount were common knowledge) in the same period. In the end of the period, he would face the choice of whether to exercise his option to buy the used unit with the strike price that was specified by the lease contract. If not, the unit would be returned to the producer and sold in the subsequent used-goods market.

The new-good market is modeled as a fixed take-it-or-leave-it lease contract. The lease term is one period. The lease price and the strike price were common knowledge and remained constant throughout each experiment. Since almost all automobile lessors have been using auction (to dealers, not consumers) as the standard method to re-market used cars, we have decided to use a round-based ascending bid auction for the used-good market. All the supply in the used-good market came from returned off-lease units. Thus, both the size and prices of the market were endogenously determined. We assume that

the residual quality of a good in the used-goods market is observable in our experiment, and thus we sidestepped the adverse selection problem that was made notorious by lemons in used-car business between individual sellers and buyers. The rationale behind our choice is based on the following two facts. First, the most relevant parts of the used-car market for auto producers are those related to off-leases or fleet rentals that are relatively new, typically one to two years old. Second, before the auction process, almost all used-cars are inspected and the results are well documented and disseminated to any potential buyers; and any deals that are in dispute can be conveniently settled through arbitrations.

The experimental observations have largely confirmed the qualitative features of the theory, both at the aggregate and individual levels. Furthermore, the comparative statics of the experimental market in responding to the change of the strike price are consistent with that of the theoretical model. On the other hand, the experimental results also provide evidence suggesting that there are systematic biases in the theoretical model due to the assumptions of perfect rationality and risk aversion.

The remainder of the paper is organized as follows. In section 2, we first recapitulate the theoretical model introduced by Huang and Yang (2002), and then briefly outline how the model is solved. In section 3, we spell out the details of the experimental design. The experimental results along with comparison with theoretical predictions are presented in section 4. We conclude and point out some future directions in section 5. The appendix provides some additional results that are relevant for the discussions in the main text.

2 The Theoretical Model

The content of this section is extracted from Huang and Yang (2002). All details can be found in the original paper.

2.1 The Goods and Lease Contract

All the goods have a lifetime of two periods. They are regarded as homogeneous when they are new. This allows us to normalize the quality measure for the entire life span of the goods to be 1. Depending on the usage of a good in its first period, the residual quality of the good in the second period is denoted by $\delta \in (0,1)$. Since the usage of a particular good is uncertain at the onset of the lease contract, δ is treated as stochastic for lessees[1] and is assumed to obey an exogenous distribution with a known density $g(\delta)$. For convenience, we take this distribution as a lognormal distribution with the following mean and volatility parameters: $\mu = -1$ and $\sigma = 0.2$. We further define $f(x) = \int_0^x g(\delta)\, d\delta$ and $\Phi(x) = \int_0^x f(\delta)\, d\delta$. On the other hand, when a used good enters the used-good market,

[1] If the residual quality were known to the lessee at the signing of the lease contract, there would have been no risk factor in each consumer's decision-making process, at least theoretically. This, in turn, would have made the option embedded in the lease contract meaningless.

δ is treated as observable for all participants there. We further assume that any remaining quality of a used good is completely consumed in the second period.

The lease contract allows the lessee to use a new good for one period with a lease price of *r*. At the lease end, the lessee has the option to either keep the used good by paying a pre-determined strike price *k* or returns the used unit to the producer without additional obligation.

2.2 Consumer Preference

From the consumers' point of view, the goods are differentiated vertically. Consumer's heterogeneity is parameterized by θ, representing the willingness to pay for a unit of quality. We assume that the distribution of θ is uniform on [0, 1] and does not change over time. In the context of the experiment, each individual will have the same θ for the whole experiment. The only uncertainty an individual can encounter is when he leases a new good: he is unsure of the residual quality δ of the leased good at the lease end.

The utility flow of an individual at each period is a function of δ and θ, as well as a function of pertinent prices: lease price *r*, strike price *k*, and used-good price $q(\delta)$ for a unit of used good with residual quality δ. To simplify the setting, we assume that the transaction cost for a consumer to sell used goods is prohibitively high. We further exclude the outright selling of new goods, in order to focus on studying the optionality embedded in the lease contract. No outright selling also implies that there is no trade-in. The following table details the assignment of the relevant utility flows for consumer θ in the case when all new goods are only leased. Since we are dealing with durable goods with transaction cost, the utility flow will have to be explicitly state dependent.

Table 1 The utility flow matrix with a state *s* and an action *a* for consumer θ: $\Pi_\theta[s,a]$

$s \setminus a$	L_δ (with unknown δ)	C (with knwon δ')	U_δ (with known δ)
$L_{\delta'}$	$(1-\delta)\theta - r$	$\delta'\theta - k$	$\delta\theta - q(\delta)$
$\bar{L}$	$(1-\delta)\theta - r$	$-\infty$	$\delta\theta - q(\delta)$

In the above table, $L_{\delta'}$ denotes a state that the consumer leased a new unit in the last period and has a known residual quality δ' entering the current period. $\bar{L}$ represents any state that the consumer did not lease a new good in the last period. L_δ (with unknown δ) depicts the action of leasing a new good in the current period with a consumed quality of 1-δ. C (with known δ') signifies the action of exercising the option to keep the used good of residual quality δ' that was leased in the last period. U_δ (with known δ) is the action of buying a used good with an observed residual quality δ. The lease price *r* and strike price *k* are announced by the producer at the beginning of every period, and are kept constant throughout the experiment.

One can easily recognize that consumers are assumed to be risk neutral in the theoretical model. While simplifying the mathematical treatment, some of the detailed quantitative discrepancy between the theory prediction and experimental observation may be attributed to the risk neutrality assumption.

2.3 Consumer Behavior: Theory

We will only be concerned with the steady limit of the dynamic equilibrium where player's reaction function becomes independent of time, and each consumer adopts a constant consumption pattern (a fixed sequence of strategies).

Concept of Solution: The dynamic aspect of the consumers' decision-making is modeled using the solution concept of Markov perfect equilibrium developed by Maskin and Tirole (1988). Strategies that a consumer can take depend only on the current state. A general equilibrium is embedded into the game at every period to endogenize the used-goods market in a style of Huang, Yang and Anderson (2000). The used-goods price is determined by the clearance condition for each realizable residual quality in the used-goods market. Grossing over the microeconomic process of price formation is again for the technical tractability. Therefore, we should expect the theory to make sense only on average, and anticipate that experimental results, which are obtained from explicitly treating the used-goods market as an ascending auction, are going to deviate from theory predictions at some detailed level. For the justification of how a competitive price emerges from auction processes, readers are referred to Wilson (1977) and Milgrom (1981).

Consumers' Bellman Equation: Given the various prices in the steady limit, consumer q at state s solves the following Bellman equation

$$V_q[s] = \max\left\{ E_d\left[\Pi_q[s, L_d] + \rho V_q[L_d]\right], \Pi_q[s, C] + \rho V_q[C], \max_{d \in \Delta_U}\left(\Pi_q[s, U_d] + \rho V_q[U_d]\right)\right\},$$

where $\rho \in [0,1]$ is the discount factor and Δ_U stands for the set of realizable residual qualities in the used-goods market. The first term in the curly brackets corresponds to leasing a new good, the second term corresponds to exercising the option, and the third term corresponds to buying a used good. That an expectation with respect to d appears only when the consumer chooses to lease reflects the fact that the residual quality d is ex ante for new leases and is ex post for actions associated with used goods.

Consumer segmentation: When the exogenous parameters of the model are appropriately chosen, the above Bellman equation admits a unique solution with the following consumer behavior. Consumers are naturally segmented by a pair of division points θ_m and θ_M (with $0<\theta_m<\theta_M<1$). Low valuation consumers, $\theta\in(0,\theta_m)$, choose to stay out of the market. Consumers in (θ_m,θ_M) choose to buy used goods. High valuation consumers, $\theta\in(\theta_M,1)$, choose to lease new goods or exercising options according to the reaction function

$$R_\theta[L_\delta]=\begin{cases}L, & \text{if } \delta<z(\theta)\\ C, & \text{if } \delta>z(\theta)\end{cases}.$$

This threshold rule leads to the following probabilities for consumer θ to be in leasing a new good or continuing to consume the used good by exercising the option: $h_L(\theta)=1/[2-f(z(\theta))]$ and $h_C(\theta)=[1+f(z(\theta))]/[2-f(z(\theta))]$. The values of the division points are determined from the conditions that consumer θ_m is indifferent in staying out of the market or buying a used good with an arbitrarily low residual quality, and that consumer θ_M is indifferent in buying the used good with the highest realizable residual quality in the used-goods market δ_M or leasing a new good.

Average Payoff Function: The average payoff function per period $(1-r)V_\theta[\cdot]\equiv\tilde{V}_\theta[\cdot]$ can be shown to have a finite limit when $r\to 1$:

$\tilde{V}_\theta[I]=0$ for $\theta\in(0,\theta_m)$;

$\tilde{V}_\theta[U_\delta]=\delta\theta-q(\delta)$ for $\theta\in(\theta_m,\theta_M)$;

and

$$\tilde{V}_\theta[L_\delta]=\tilde{V}_\theta[C]=\frac{[1+\Phi(z(\theta))-z(\theta)f(z(\theta))]\theta-[1-f(z(\theta))]k-r}{2-f(z(\theta))},\qquad \theta\in(\theta_M,1).$$

As it is well-known, the $r\to 1$ limit is called time-average criterion. We will justify later why this criterion is the relevant one for the experimental setting.

Option-exercising Threshold: The option exercising threshold is related to the average payoff function as $z(\theta)=(k+\tilde{V}_\theta[C])/\theta$. In the limit of $r\to 1$ the threshold satisfies the simple equation:

$r-k=[1+\Phi(z(\theta))-2z(\theta)]\theta$.

Used-good Market: The used-goods supply is from off-leases. The price for a used good with residual quality δ is determined by the clearance condition:

$$\theta(\delta)-\theta_m=\int_{\theta_M}^{1}d\theta\,\frac{f(\min\{\delta,z(\theta)\})}{2-f(z(\theta))}.$$

Supplementing the clearance condition with the equation of marginal substitution rate $\theta(\delta)=\frac{dq(\delta)}{d\delta}$ and the terminal condition $q(0)=0$, the price for a used good with residual quality $\delta\in(0,\delta_M)$ can be written as

$$q(\delta) = \theta_m \delta + \int_0^{\delta} d\delta' \int_{\theta_M}^{1} d\theta \; \frac{f(\min\{\delta', z(\theta)\})}{2 - f(z(\theta))}.$$

Residual Quality Distribution in Used-goods Market: The residual quality distribution in the used-goods market is modified from the original residual quality distribution according to

$$\tilde{g}(\delta) = g(\delta) \int_{\theta_M}^{1} d\theta \; \frac{I[\delta < z(\theta)]}{2 - f(z(\theta))} \Bigg/ \int_{\theta_M}^{1} d\theta \; \frac{1}{2 - f(z(\theta))},$$

where $I[\cdot]$ is the indicator function. The return rate is obtained by integrating over δ,

$$w(r,k) = \int_0^1 d\delta \; \tilde{g}(\delta) = \int_{\theta_M}^{1} d\theta \; \frac{f(z(\theta))}{2 - f(z(\theta))} \Bigg/ \int_{\theta_M}^{1} d\theta \; \frac{1}{2 - f(z(\theta))}.$$

Numerical Solution: Some of the equations, such as the clearance condition and threshold equation, do not appear to be amenable in closed form. However, they can be easily solved numerically.

3 Experimental Design

The experimental model was implemented in the HP Experimental Economics Software. Every experiment has around twenty five subjects each playing the role of a consumer who procures a durable good that lives for two periods. Each person is limited to process at most one unit of good in any given period. Instructions for the experiments were posted on the web. Each subject had to pass a web-based quiz before he was allowed to participate. The instructions and the accompanying quiz are available at:
http://www.hpl.hp.com/econexperiment/lease/instructions.htm

3.1 Preference

Preferences were induced according to the vertical differentiation model described above. The homogeneity of the new goods means that we can normalize the quality measure for the entire life span to be 1. The residual quality for a used unit is denoted by $\delta \in (0,1)$. This parameter divides the whole quality into new and used. δ is drawn from a known distribution.

Consumer's heterogeneity is parameterized by $\theta \in [0,1]$, which represents the willingness to pay for a unit of quality. The theoretical analysis assumes that θ is drawn from a uniform distribution and that θ does not change over time. The experimental design deviates slightly from these assumptions. We chose θ s to span over the [0,1] interval in the following manner. Consider an experiment with N subjects. The interval [0,1] was divided into N equal intervals: [0,1/N], (1/N,2/N], … ((N-1)/N, N]. Each of these "mini" intervals was then assigned to a different subject. Each subject's θ was drawn randomly from his "mini" interval. This design ensured we would observe θ s to span uniformly over the [0,1] interval.

Each subject was allowed to switch q once in each experiment. This was done usually on period 13. If a subject drew his first q from the interval (m/N, (m+1)/N], his second q would be drawn from the interval ((N-m-1)/N, (N-m)/N]. Thus, if a subject had a very low first q, his second q would be guaranteed to be high. This was done to address fairness concerns.

Furthermore, each subject was given $1 for each period he completed since only a subset of the subjects were expected to make money.

3.2 Decisions

In each period, an individual has the following choices of action: 1) hold no unit; 2) lease a new unit; 3) purchase the leased unit with the strike price k if at a lease-end; and 4) buy a used unit from auctions held by the producer.

New units were leased from an exogenous source, referred to as the producer in the experiment. The lease price and the strike price were announced to the subjects in the beginning of each period and they understood that these prices were common knowledge. In all the experiments, both the least price and lease-end strike price stayed the same throughout the experiment. The lease price was paid at the onset of the lease contract and entitles the lessee to use a new good for one period. At the beginning of the 2^{nd} period of the life of the leased good, the individual has two alternatives. He could purchase this unit at the strike price or he could decide to return the unit to the producer. The residual quality is unknown at the time of signing the lease contract and becomes known at the time of deciding whether to exercise the option. If a unit was returned to the producer, it would be sold in an auction in the following period.

Simultaneous round-based ascending bid auction was chosen as the auction mechanism. This is similar to the actual used car auctions that most of the auto producers are conducting, except that in our case the units are directly sold to the consumers while large scale used-car auctions are only open to dealers. Since auction was also well studied in the literature (Kagel et al. 1995), this choice also enables us to interpret our results in the context of past experiments if the need arises. Residual qualities were announced before each unit was auctioned. Furthermore, subjects did not have to decide on a new lease or exercising of an option until the end of the auction.

3.4 Treatments

The key issue of importance to the producer's used-car remarketing business is how the endogenous used-car market reacts to changes in the lease contracts. In particular, the producer is interested in the effect of the lease-end strike price. Two treatments were used in the experiments, one with a strike price equal to k=0.08 and the other with a strike price of k=0.16.

4 Results

4.1 Overview

As outlined in the Introduction, there are dual motivations for carrying out the experimental study. The first is to check whether the economic assumptions adopted by the theoretical model, such as perfect rationality and risk neutrality, are plausible. The second is to gauge the robustness of the theoretical predictions on the price formation mechanisms in the used-good market given that not all the assumptions of the model would hold true when real human beings are involved. Due to its lack of an explicit modeling of the microeconomic process in the used-good market, we can only hope that the theoretical model makes quantitative sense on average at best. In addition, the number of subjects is about 25 for each experiment, which may not appear to be small from the first sight. However, taking into account the consumer's heterogeneity and the complexity of possible consumption decisions that can endogenously emerge; we still expect substantial finite-sample fluctuations. Therefore, when contrasting the experimental results with their theoretical counterparts, we will mostly concentrate on qualitative and comparative static aspects.

For the sake of convenience, all valuations and prices in the experiment are rescaled by a factor of 1000, so that subjects can submit bids that are integers. A total of 4 experiments were conducted. The following table provides an overview of all the experiments.

Table 2 Summary of the four experiments.

Experiment	Number of subjects	Number of periods	Lease price	Strike price	Number of new leases	Number of auction units
1	25	25	300	80	209	54
2	28	24	300	160	158	85
3	28	23	300	160	174	122
4	23	21	300	160	151	95

Ideally, more experiments would be conducted. However, business constraints only allowed for 4 experiments in this study. Despite the small number, experimental results seem to be robust with respect to some qualitative features of the model.

Two issues arose in the course of this work that resulted in the choice of parameters (three experiments with strike price of 160). The first issue is sampling effects. Typical experiments at HP Labs use Stanford students who already had prior experience through participating in other earlier experimental economics projects. Around the same time when Experiment 2 was conducted, HP had expanded its scope of subject recruitment to a local city college, mainly due to the need of a completely different project. As a result, some students from the local city college were conveniently recruited for Experiment 2. These students from the local city college never had any prior experience. Unfortunately, this fact was overlooked during the process of training subjects to become proficient in decision-making in the context of this experiment. As a consequence, significant portion

of the subjects earned substantially less than predicted, whereas this was not true for the rest of the subjects in the same experiment or in other experiments (see Appendix). These subjects had a strong inclination to participate and to win auctions even when their best option was to choose other strategies.

A second issue was a good illustration of how the detail in market institution design matters. The rule of the game in the first three experiments was that a player could choose either to bid in an auction *or* to start a new lease but not both. In effect, at the last round of auction, a player has to choose between bidding, which could result in no win, and start a new lease. Consequently, some high valuation consumers, who decided to participate in the auction process but were not able to win the bid in the final round, were deprived of the chance to lease new goods. Obviously, the theory has no bearing on this issue since it is based on a market clearing and does not specify an explicit mechanism of how the clearance is achieved. This issue turns out to be not severe for Experiment 1, because the size of the used-good market is small, but it is quite noticeable in Experiments 2 and 3.

These two issues compounded together resulted in fewer new leases in Experiments 2 and 3, in which some of the high valuation subjects achieving very low payoffs. This prompted us to change the game rule in Experiment 4 from "either to bid or to lease new" to that a subject can always have a chance to lease new if he loses in the auction. We believe that the latter rule is closer to reality. Thus, we conducted 3 experiments using k=160 with slightly different market rules and different samples of subjects. Ideally, we would conduct more experiments to contrast the effects of these issues. However, the rigorous time table of a business related project did not allow us to do that. Furthermore, while this is not standard experimental methodology, we do not believe these variations substantially altered our conclusions.

Before we present the details, it is important to emphasize that there are no free parameters in the theoretical predictions when they are compared with their corresponding experimental counterparts, once the exogenous parameters, such as the distributions of consumer's heterogeneity and residual qualities, are chosen to be the same. However, finite sampling sizes in the experiment can introduce systematic bias to these distributions. To partially alleviate this kind of bias, especially for the distribution of residual qualities, the theoretical predictions are calculated based on the values of parameters computed directly using the finite samples realized in the experiments. Due to our way of sampling q, finite sampling effect for consumer heterogeneity is less of a problem. Finally, we fix the value of the only behavior parameter in the theory as $r \to 1$, or according to the time-average criterion. This choice can be justified by the fact that subjects' monetary rewards are mostly based on their cumulative performance in more than 20 periods. Thus, the experimental setting is such that subjects are motivated to maximize their average payoff per period.

4.2 Aggregate Level Comparison

Among the aggregate variables that we examine are the following: 1) new-lease probability per period per consumer, which serves as a measure for the demand of new goods; 2) return rate, which measures how likely a lessee exercises the embedded option in the lease contract; 3) average used-good price, which is endogenously determined; 4) aggregate surplus per period, which measures how consumers as a whole benefit from participating in the market; and 5) producer revenue per period, whose sources of contribution include new-leases, exercised options and resale of used goods. Variables normalized by the number of periods and/or number of subjects will enable us to combine results obtained from different experiments of the same setting, and to compare results from different experimental settings in a meaningful manner.

In order to have an appreciation of how finite sampling correction affects the theoretical prediction, we first list these predictions with the originally chosen parameters for the residual quality distribution $\mu = -1$ and $\sigma = 0.2$ in Table 3. Typically, the finite sampling implies about 5% corrections to the mean and 10% corrections to the volatility. As we will see shortly, all aggregate variables, except return rate, are not very sensitive to the finite sampling correction.

Table 3 Theoretical predictions with the intended parameters of the residual quality distribution: $\mu = -1$ and $\sigma = 0.2$.

Theory	New-lease prob. per period per consumer	Return rate per lease	Average used-good price	Aggregate surplus per period per consumer	Producer revenue per period per consumer
k=80	0.34	0.53	105	104	135
k=160	0.33	0.89	126	96	144

Table 4 lists the results of Experiments 1 to 4, along with the corresponding theoretical predictions corrected by the finite-sampling effect. Since Experiments 2, 3 and 4 share the same k=160, we first average the aggregate results from these three experiments and then compare the average to the theory. The differences between these three experiments also serve as a crude measure of behavior fluctuations from rather small sample sizes of subjects. Given the fact that there is no fitting process involved in the comparison, the level of the agreement between experimental results and theoretical predictions in Table 4 is quite remarkable. Quantitatively, the worst case is the return rate, in which the experimental values are systematically lower than that of the theory by about 30%. One way to interpret this systematic difference is risk aversion. The only uncertainty in this model is the consumption in the first period of a new lease, represented by an unknown residual quality that is only realized at the lease-end. Thus, risk averse agents may be inclined to keep the leased unit, whose value is known at the time of exercising the option, instead of starting another new lease. Consequently, return rate will be lower than the theory that assumes risk neutral consumers. Another possible way to interpret the systematic discrepancy may be traced to ownership effects. However, to settle the true cause, additional theoretical modeling and experimental investigation are needed.

Table 4 Experimental results and theoretical predictions with the finite-sample parameters of the residual quality distribution realized in each experiment.

Experiment	New-lease prob. per period per consumer	Return rate per lease	Average used-good price	Aggregate surplus per period per consumer	Producer surplus per period per consumer
1 (k= 80)	**0.33**	**0.26**	**115**	**94**	**130**
Theory [2]	*0.33*	*0.37*	*113*	*101*	*130*
2 (k=160)	0.24	0.54	147	52	107
3 (k=160)	0.27	0.70	122	70	117
4 (k=160)	0.31	0.63	90	88	130
Average(2,3,4) (k=160)	**0.27**	**0.62**	**120**	**70**	**118**
Theory [3]	*0.32*	*0.80*	*132*	*91*	*142*

A primary policy question that a producer is interested in is how the market would respond to a change in the strike price. The theory predicts that an increase in the strike price from k=80 to k=160 at a fixed lease price will lead to a slight decrease in total lease volume, a substantial increase in the return rate, an increase in average used-good price, a reduced aggregate surplus for consumers, and an increase in producer revenue. All these directional changes are confirmed in Table 4, with the exception of producer revenue, which went the opposite way of the theoretical prediction. We attribute this deviation to the fact that there are too few new leases in Experiments 2 and 3, caused by issues of market rules and subject sampling mentioned earlier. It is worth noting that the theory predicted a substantial change only in the return rate while all other changes are more moderate. Experimental results confirmed this substantial change in the return rate.

We chose not to report standard deviation statistics. Since the game is dynamic in nature, data across periods were not independent. Thus, calculating standard deviations, or any other variance estimates, across periods would not be useful. Furthermore, variations in subject behavior were mostly driven by their different willingness-to-pay parameter θ. Therefore, reporting variance estimates across individuals would not truly reveal heterogeneous individual characteristics such as risk aversion. However, most of the comparative static holds true between any of Experiment 2, 3, or 4 (with k=160) and Experiment 1 (with k=80). Thus, we have some confidence that the comparison is valid.

4.3 Detailed Level Comparison

We now examine how the experimental results and theoretical predictions compare at a detailed level. In particular, we are interested in seeing how patterns of consumer behavior emerge as a function of willingness-to-pay. We are also interested in seeing how used-good prices change with variations of residual quality. For the sake of space limitation, we will only use the results for Experiment 1 as illustrating examples. In most

[2] The finite-sample parameters of the residual quality distribution realized in Experiment 1 are $\mu = -0.95$ and $\sigma = 0.18$.

[3] The finite-sample parameters of the residual quality distribution realized in Experiments 2, 3 and 4 are $\mu = -0.96$ and $\sigma = 0.22$.

of the cases, the results of Experiment 1 are quite typical. Due to the fact that the used-good market is treated tersely in the theory, we expect that the theory will fare less well at a detailed level than at an aggregate level.

In the following we treat the same subjects with a different q essentially as a different consumer. If all the data were used, each subject would yield two points. Thus, we observe a total of twice as many consumers as the number of subjects in each experiment. It can be argued that the data in the first two periods with freshly assigned q values should be thrown away because of start-game effects. However, we found that the conclusions are not dependent on whether we exercise this option.

4.3.1 Average Payoff and Used-good Price

Figure 1 shows average payoff per period as a function of consumer heterogeneity q. In the left panel of the figure, the theoretical payoff curve tracks very closely the experimental payoffs. The right panel of the figure indicates that the observed used-good prices are clustered around the theoretical prediction. The trend that higher residual quality implies a higher used-good price is reproduced, though with large fluctuations. There is a small number of observations whose residual qualities are higher than the point where the theory curve ends. This signals a slight behavior deviation from the theory, which predicts that there is an upper limit in residual qualities in the used-good market due to the presence of the option. Nevertheless, Figure 1 allows us to conclude safely the following results.

<u>Result 1: Observed payoffs are consistent with the theory.</u>

<u>Result 2: Observed used-good prices are consistent with the theory.</u>

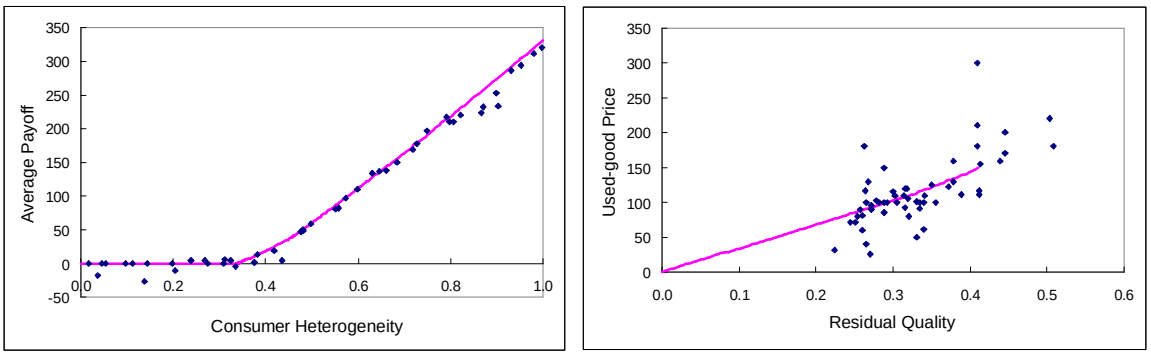

Figure 1 Average payoff as a function of consumer heterogeneity (left panel) and used-good price as a function of residual quality (right panel). Curves are theoretical predictions and diamond points are experimental observations in Experiment 1.

4.3.2 Behavioral Segmentation

The theoretical model predicts that subjects would be segmented endogenously into three classes of behavior. Lower valuation consumers $q \in (0, q_m)$ are priced out of the market.

Medium valuation consumers in $\theta \in (\theta_m, \theta_M)$ participate in the used-good market. High valuation consumers $\theta \in (\theta_M, 1)$ lease new goods and occasionally exercise the option embedded in the lease contract at lease-end.

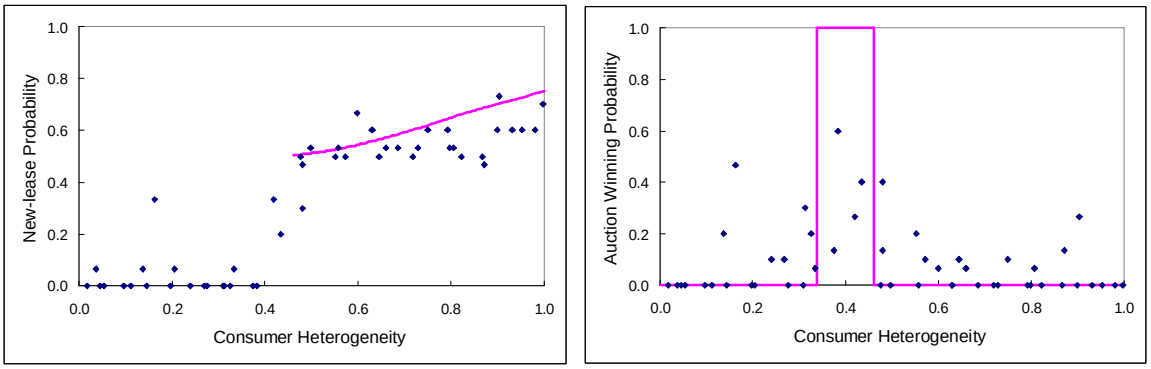

Figure 2 New-lease probability (left panel) and auction winning probability (right panel) as functions of consumer's heterogeneity. Lines are theoretical predictions and diamond points are experimental observations in Experiment 1.

Behavior segmentation can be captured in two measures: new-lease probability and auction-winning probability. Figure 2 shows these probabilities as functions of θ. In Experiment 1, the theory predicts $\theta_m = 0.33$ and $\theta_M = 0.47$, respectively. As one can see from Figure 2, both new-lease probabilities and auction winning probabilities are quite low when θ <0.3. This supports the conclusion that on average, low valuation consumers are priced out of the market. New lease probabilities begin to rise at around θ =0.4 and become quite close to the theoretical curve from around θ =0.5 onward. On the other hand, though still roughly concentrating at around the right region, auction-winning probabilities are much more spread than the theory's prediction. From time to time, consumers who would be theoretically the pure used-good buyers also enter the new-lease market, and consumers who would be theoretically pure lessees venture into the used market. One interpretation is that the fundamental economics forces were operating correctly. However, the perfect rationality assumption in the theory is obviously violated, leading to the smearing in consumer segmentation.

Interestingly, the smeared behavior does not cause a substantial payoff gap, as can be inferred from the left panel in Figure 1. This implies that the economic incentive that is responsible for the sharp segmentation in theory is not very strong for those consumers whose willingness-to-pays are in the middle, and occasional "mistakes" are gracefully tolerated. In addition, Figure 2 also provides evidence on why several subjects have their payoffs much lower than the theoretical curve. For example, consumers whose θ values lie between 0.8 and 0.9 should have leased more new goods rather than participated in auctions. Nevertheless, the following conclusion can be drawn.

Result 3: Strong but Imperfect Patterns of Behavioral Segmentation.

4.3.3 Cherry Picking

Theoretically, units with a higher residual quality have a higher chance of being purchased by the consumer exercising his lease-end option. Thus, the units returned to the producer would have a distribution skewed towards the low-end compared to the original distribution of residual qualities. This kind of *cherry picking* phenomenon is also observed in the experiment. Figure 3 shows the distribution of residual qualities for all the units and the distribution for those units that were returned to the producer and subsequently entered the used-good market. Notice that not all high residual quality units were returned to the producer as predicted.

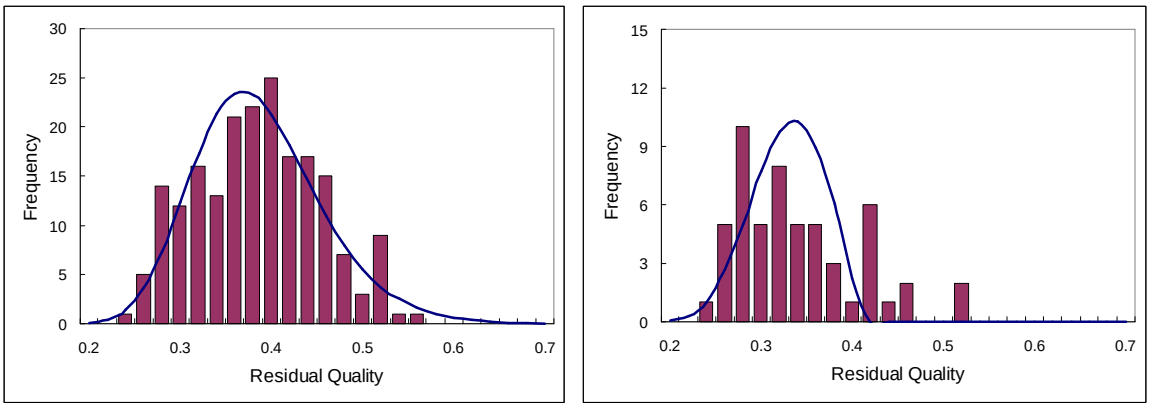

Figure 3 Distributions of residual qualities for all used units (left panel) and for those that enter the used-good market (right panel). Bars are experimental observations in Experiment 1, and curves are theoretical predictions, which are normalized to have the same masses as in the experiment.

Furthermore, Kolmogorov-Smirnov Tests (Table 5) show that, in three out of four experiments, the distribution of residual qualities of the returned units is consistent with model predictions. Experimental evidence not only confirms the cherry picking phenomenon in a *qualitative* fashion, but also suggests that the theory is sound *quantitatively* despite all the handicapping factors mentioned before.

Table 5 Kolmogorov-Smirnov Test to see if residual qualities of the returned units were consistent with the theoretical distributions.

Experiment	Observations	K-S Statistics	P-Value
1	54	0.177	0.97
2	85	0.092	0.78*
3	122	0.069	0.70*
4	95	0.057	0.48*

* cannot reject the null hypothesis at 95% confidence that observed residual qualities obey the distribution specified by the theoretical model.

Result 4: Cherry Picking Observed and Consistent with Theory.

5 Conclusion

A sequence of experiments was conducted at Hewlett-Packard Labs, in collaboration with Ford Research Lab, to study consumer behavior in a durable goods market where leasing is prevalent. The experiments have mostly confirmed aggregate predictions of the theory

and validated several qualitative features of the theoretical model. We observed subjects segmenting themselves into classes of behavior based on their willingness-to-pay parameters. Subjects at the low end of willingness-to-pay were priced out of both the used- and the new-goods markets. Subjects at the high end leased with increasing frequencies. They sometimes exercised their options depending on the realization of the residual quality and the potential value achievable at the used-goods market. The last segment of the subjects lived in the middle and primarily participated in the used-goods market. The sizes of these three groups were qualitatively consistent with the theoretical model. Furthermore, when we increased the strike price in a different treatment, the experimental market mostly responded in the direction predicted by the model. This result is robust even with small variations of market rules and sampling of subjects. Given the fact that the theoretical model has largely grossed over issues of market rules in the used-good market, the near agreement between the theory and experiment is highly non-trivial.

On the other hand, in all the experiments, the subjects with high valuation are more likely to exercise the option relative to the theoretical prediction. There are multiple possible explanations. One such possibility is risk aversion that is not addressed by the theoretical model. With risk aversion, a leasing subject has the tendency to keep the used unit that entails no uncertainty relative to lease a new good that has an unknown consumption in the first period. Other explanations such as ownership effects may also account for the discrepancy between theory and experimental results. More evidence is needed to pinpoint the correct explanation.

The effect of learning in the experiment appears to manifest mostly in whether subjects are used to the economic context of the experiment. Once subjects familiarize themselves with the decision-making process, there is no obviously discernable effect associated with progressive stages of the experiment. However, due to the complex setting of the experiment, less experienced subjects, as exemplified in Experiment 2, took a long time to figure out what they ought to behave and hence earned significantly less payoffs comparing to more experienced subjects.

There are several directions that can be viewed as natural extensions of the current work. To settle whether the aforementioned systematic bias in return rate is caused by risk aversion or something else can be pursued by extending the theory to include risk aversion and conducting additional experiments that are specifically designed for this purpose. Another interesting direction is to treat the residual quality being only partially observable, which in turn will allow the possibility of studying the interplay between optionality and adverse selection. Investigations of lease contracts with more sophisticated options and under oligopoly market structure are other topics for future exploration. In addition, it is important to realize that the setting of the current experiment is not very far from many realistic business environments. Adapting the experiment described in this paper to field studies has the potential to provide useful business insights. Finally, work has already begun to use a modified version of this experiment to examine business strategies in other aspects of the automotive market.

Appendix: Payoff Curves

The following figures show payoffs of all four experiments. Each point represents the average payoff of a subject under the same willingness-to-pay parameter. It is interesting to note that earnings for all subjects in Experiments 1 and 4 and for most subjects in the other two experiments are very close to the predicted values. As pointed out in section 4, some subjects in Experiments 2 and 3 were earning substantially less money.

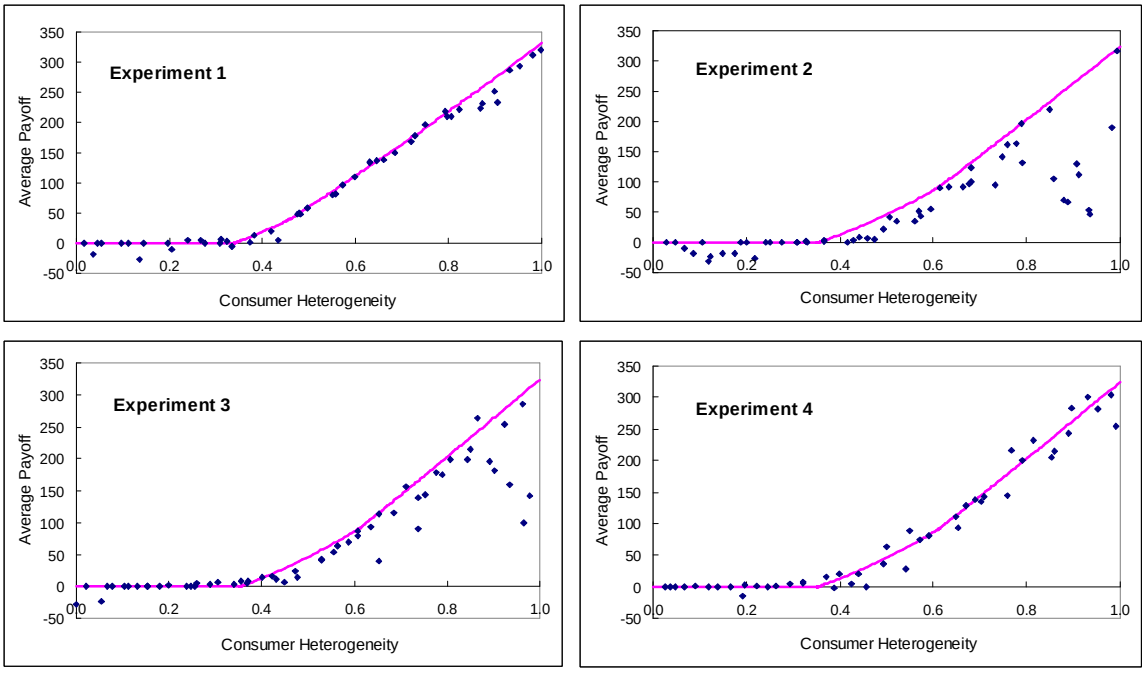

Figure 4 Average payoff as a function of consumer heterogeneity in all four experiments.

References

Black, F. and Scholes, M., 1973, "The Pricing of Options and Corporate Liabilities", *Journal of Political Economy*, **81**, 637-659.

Camerer, C. "Individual Decision Making", In *The Handbook of Experimental Economics*, edited by Kagel J. and Roth A. Princeton Univ. Press. 1995.

Huang, S. and Yang, Y., 2002, "Pricing Lease Contracts with Options in Imperfect Markets of Durable Goods", Technical Report, Ford Research Laboratory.

Huang, S. Yang, Y. and Anderson, K., 2001, "A Theory of Finitely Durable-Goods Monopoly with Used-Goods Market and Transaction Costs", *Management Science*, **56**, 549-569.

Kagel, J. “Auctions: A Survey of Experimental Research”, In *The Handbook of Experimental Economics,* edited by Kagel J. and Roth A. Princeton Univ. Press. 1995.

Maskin, E. and Tirole, J., 1988, "A Theory of Dynamic Oligopoly: I and II", *Econometrica*, **56**, 549-569 and 571-599.

Merton, R., 1973, "Theory of Rational Option Pricing", *Bell Journal of Economics*, **4**, 141-183.

Milgrom, P., 1981, "Rational Expectations, Information Acquisition, and Competitive Bidding", *Econometrica*, **49**, 921-943.

Wilson, R., 1977, "A Bidding Model of Perfect Competition", *Review of Economic Studies*, **44**, 511-518.