Is econometrics useful for *private* policy making? A case study of replacement policy at an auto rental company

Sungjin Cho [a,1], John Rust [b,*]

[a] *Seoul National University, Republic of Korea*
[b] *University of Maryland, United States*

ABSTRACT

The goal of this paper is to illustrate the *potential* usefulness of econometrics as a tool to assist *private* policy makers. We provide a case study and detailed econometric analysis of the automobile replacement policy adopted by a large car rental company. Unlike public policy making – where the benefits from using econometric models and "science-based" approaches to policy making are hard to quantify because the outcomes of interest are typically subjective quantities such as "social welfare" – in the case of firms there is an objective, easily quantifiable criterion for judging whether policy A is better than policy B: *profits.* We introduce and estimate an econometric model of the rental histories of individual cars in the company's fleet. Via stochastic simulations, we show that the model provides a good approximation to the company's actual operations. In particular, the econometric model is able to reproduce the extraordinarily high rates of return that the company obtains on its rental cars, with average internal rates of return between purchase and sale of approximately 50%. However, the econometric model can simulate outcomes under a range of counterfactual vehicle replacement policies. We use the econometric model to simulate the profitability of an alternative replacement policy under pessimistic assumptions about the rate maintenance costs would increase and rental rates would have to be decreased if the company were to keep its rental cars longer than it does under the *status quo.* Depending on the vehicle type, we find that the company's expected discounted profits would be between 6% to over 140% higher under the suggested alternative operating strategy where vehicles are kept longer and rental rates of older vehicles are discounted to induce customers to rent them. The company found this analysis to be sufficiently convincing that it undertook an experiment to verify the predictions of the econometric model.

© 2008 Elsevier B.V. All rights reserved.

Executive summary

This paper illustrates the use of econometrics as a tool to improve policies chosen by *private* policy makers. We argue that one of the most convincing demonstrations of the value of econometric modeling is to show how it can be used to help firms find more profitable operating policies. This point is illustrated with a case study of the vehicle replacement policy of a large rental car company.

Most major rental car companies (e.g. Hertz, National, etc.) typically sell cars after only 20,000 miles. However, this is an extremely costly policy due to the well known rapid depreciation in car prices. We study the profitability of the "rapid replacement policy" using an econometric model of the operations of a specific rental car company. This model is sufficiently detailed that it can account for individual rental contracts (including short term rental and long term rentals, i.e. monthly lease contracts), and durations of cars on the lot waiting to be rented. In addition to this multistate semi-Markov model of the car's contract status, we also estimate submodels that predict maintenance costs and the secondary market (resale) price the company receives when it decides to sell their cars. The model is found to be a remarkably accurate description of the overall operations of this firm with simulations of the model accurately predicting a wide range of operating and financial outcome variables of different makes and models of vehicles in the company's fleet.

Our econometric model predicts that the company can significantly increase its profits by keeping its rental cars longer than it currently does — even under pessimistic or "worst case" assumptions. We find that profits can be significantly increased if vehicles are replaced after roughly 90,000 miles (or about 5 years) instead of 3 years (or about 45,000 miles) under the company's

* Corresponding address: Department of Economics, University of Maryland, 3105 Tydings Hall, College Park, MD 20742, United States. Tel.: +1 301 404 3489.
   E-mail addresses: sungcho@snu.ac.kr (S. Cho), jrust@gemini.econ.umd.edu (J. Rust).
   [1] Department of Economics, Seoul National University, Seoul, Republic of Korea. Tel.: +82 2 880 6371.

current replacement policy. Depending on the vehicle type, the simulations predict that expected discounted profits would be 6% to over 140% higher, depending on the type of vehicle.

The company that we studied found our analysis to be sufficiently convincing that it decided to undertake an experiment to test our prediction that keeping cars longer will significantly increase its profits. The results from the experiment found that profits were in fact increased. Thus, this paper provides a concrete example of a how an econometric model can improve policy making.

In this case, the ability to demonstrate that a suggested counterfactual policy really is "better" is a result of the fact that the policymaker agrees on the objective function – maximizing profits – and the fact that "profits" is an objectively measurable quantity. But our study also illustrates another important point: to be taken seriously, an econometric model needs to be *credible*. If our model was not credible, the company would have never been persuaded to undertake an experiment to verify the predictions of the model in the first place.

Ultimately, the most important test of the credibility and usefulness of an econometric model is not just whether policy makers will undertake the data gathering necessary to build them and the experiments necessary to test them, but whether policy makers will actually be guided and change their policies in response to the feedback provided by these models. While the experiment undertaken by the company in response to the predictions of our model is an encouraging first step, it remains to be seen whether our work will really change the operating polices of the rental car company we studied.

1. Introduction

The goal of this paper is to illustrate the *potential* usefulness of econometrics as a tool to improve policies chosen by *private* policy makers. We provide a case study and detailed econometric analysis of the vehicle replacement policy of a large rental car company. Contrast this with public policy making where the benefits from using econometric models and "science-based" approaches to decision making are hard to quantify because the outcomes of interest are typically subjective quantities such as "social welfare". However in the case of firms there is an objective, easily quantifiable criterion for judging whether policy A is better than policy B: *profits.* Thus, one way to demonstrate the value of econometric models is to use them to help firms increase their profits.

In a related article, Rust (2007) presents a separate case study that focuses on whether econometrics is useful for improving *public* policymaking. That paper draws on Rust's experience as a former consultant to the US Social Security Administration. Unfortunately in that application, which involved predicting the welfare and cost/benefit impact of a proposed change in the U.S. Disability Insurance system, Rust concludes that the prospect that econometric models will have any measurable role or impact on public policymaking is bleak — at least in the short term under the current regime in Washington, DC where "faith based" policymaking and political considerations typically dominate scientific advice. In our view, the best chance to interest public policymakers in the value of econometric models in the long run is to start by providing clear-cut demonstrations of the usefulness of econometric models in improving policymaking in private sector applications in the short run. While the issue of rental policy is admittedly less interesting, ambitious, or important *per se* than Social Security policy, it is also far less politically sensitive. Also, because firms often have stronger incentives to improve outcomes and have more control over their own actions than do government bureaucrats, there is a greater chance that private policymakers

will pay attention and actually take concrete actions in response to predictions and policy recommendations from econometric models than generally is the case in public sector applications.

We illustrate these points using a new data set on rental car histories that we obtained from a major rental car company. These data provide a unique opportunity to test the hypothesis that this company has adopted a profit maximizing vehicle replacement policy. We are not aware of previous studies that have questioned the widely adopted policy by rental car companies of replacing their cars frequently. Indeed many major rental car companies (e.g. Hertz, National, etc.) typically sell cars after only 20,000 miles. However, this is an extremely costly policy due to the well known rapid depreciation in car prices.

The econometric model predicts that the company can significantly increase its profits by keeping its rental cars longer than it currently does – even under pessimistic or "worst case" assumptions to be described below. We build on results from Cho and Rust (2008) who used numerical dynamic programming to calculate optimal replacement policies under these worst case assumptions. Their findings indicate that profits can be significantly increased if vehicles are replaced after roughly 90,000 miles (or about 5 years) instead of 3 years (or about 45,000 miles) under the company's current replacement policy. We use their econometric model to conduct stochastic simulations of the company's profits under this suggested counterfactual replacement policy. Depending on the vehicle type, the simulations predict that expected discounted profits would be 6% to over 140% higher, depending on the type of vehicle. These predictions were made under pessimistic assumptions about the degree maintenance costs will increase and the amount rental rates of older vehicles would have to be decreased (to induce customers to rent them) if the company were to keep its vehicles significantly longer than it currently does. The company found this analysis to be sufficiently convincing that it decided to undertake an experiment to verify our prediction that keeping its cars longer would significantly increase its profits.

This study also illustrates the limitations of simple "reduced-form" econometric approaches to testing the profitability of the company's current rental car replacement policy, similar to the "freakonomics" approach adopted by Levitt (2006). The main value added from constructing a more explicit econometric model of the rental histories of individual cars in the company's fleet is that a wide range of outcomes can be *simulated* under both the *status quo* and a range of counterfactual replacement policies. This is not possible under Levitt's approach. Our simulations demonstrate that the econometric model provides good approximation to the company's actual operations — at least under the *status quo.* In particular, simulations of the econometric model are able to reproduce the extraordinarily high rates of return that the company obtains on its rental cars, with average internal rates of return between purchase and sale of approximately 50% under the company's current replacement policy.

The methods used in this paper are an extension and specialization of work on optimal replacement of durable goods (Rust (1985, 1987) and Cho (2008)). The main difference between the approach taken in this paper and this previous work is that these previous studies *assumed* that firms are behaving optimally, whereas in this study we relax and empirically test this assumption. It is possible for us to relax and test the maintained assumption of profit maximizing behavior in this case precisely because the firm's objective, *expected profit maximization,* does not involve subjective quantities that need to be estimated. This stands in stark contrast to most political or public policy applications, where the behavior of individuals or organizations depends critically on what their objectives are. Economists typically presume that these agents are acting "as if" they were maximizing the expected value of well-defined utility function

or a social welfare function. However, unlike profits, utility and social welfare functions are typically not known explicitly, even by the decision makers in question. As a result, strong assumptions have to be imposed in order to estimate them, including the hypothesis of *expected utility maximization.* Rust (1994) has shown that the hypothesis of expected utility maximization *per se* is not sufficient to uniquely identify an agent's underlying utility function. Indeed, there is a fundamental identification problem: it is often possible to "rationalize" many types of behavior as being optimal for some appropriately defined objective function. The identification problem is another reason why it is so difficult to apply econometric models to help improve public policymaking, because it is very hard to be sure what the "true objectives" of citizens and policymakers really are, and given incomplete knowledge of payoffs/rewards and how to aggregate and resolve conflicting objectives of the many different individuals who are affected by public policy decisions, it is hard to provide an objective basis for judging whether policy A is "better" than policy B.

Rental car companies are an ideal testing ground for an econometric approach to policymaking since these companies have large fleets, make frequent replacement decisions, and have good records of their operations and decisions. Furthermore, there is little disagreement that their basic objective is to maximize expected discounted profits, and it is is straightforward to measure profits in this application. We feel extremely fortunate to have been able to earn the trust of a large rental car company, which has provided their operational data to us, and allowed us to study and dialog with them. In order to protect the confidentiality of this firm and continue our relationship with them, we cannot provide any further information about the firm or its operations or exchange the data we obtained from the company with other researchers. However, we would like to make clear that we have no financial relationship with this company and have not received any monetary compensation for our services. In effect, we are *consulting for data* and the fact that we are doing this work with a scientific rather than financial objective in mind enables us to conduct a more independent and unbiased analysis of the company's operations.

Section 2 describes the rental car data and Section 3 introduces and estimates an econometric model of the rental company's operations. In Section 4 we show, via stochastic simulations, that this model provides a good approximation to actual outcomes for this company. Section 5 shows how the econometric model can be used to evaluate counterfactual operating strategies. We show that an alternative replacement policy of keeping rental cars longer results in significantly higher profits. We make some concluding remarks in Section 6.

2. Data and preliminary regression analysis

We obtained data from a large syndicated rental car company that owns and rents a large fleet of rental vehicles. The company provided us with data on over 3900 individual vehicles at various rental locations. These do not represent the entire fleet at any point in time, but they do represent a significant share of the company's holdings. All of these vehicles were first acquired (i.e. registered) after 1999, and almost all of these vehicles were purchased brand new from auto manufacturers. While there are occasional "group purchases" of particular brands and models of vehicles on the same date, when these group purchases did occur, they typically amounted to only 4 or 5 vehicles of the same brand/model at the same time. Thus, this company by in large follows an *individual* vehicle replacement and acquisition strategy, as opposed to "block acquisitions and replacements" i.e. simultaneously acquiring and disposing of large groups of vehicles of the same make and model at the same time.

The data include the date and purchase price for each vehicle the company acquired, the date and odometer value when the vehicle was sold, and the complete history of maintenance and rentals between the purchase and sale dates. The rental contract data record the dates each contract started and ended, and (sometimes) the odometer value of the vehicle at the start and end of the rental contract. We found (with the exception of the odometer value at the date each vehicle was sold, which was accurately recorded), the company's data on odometer values at the beginning and end of each rental contract to be frequently missing or based on guesses by the company's rental agents. This was especially true for long term contracts that were "rolled over" (i.e. where the contract was renewed by the customer without returning it back to the lot). Many of the company's rental agents appear to have filled in rough estimates of the out and in odometer values at the roll over dates when customers informed them of their to keep their car another month. As a result, we did not trust most of the in or out odometer readings in the company's rental records. In order to infer the driving patterns and number of kilometers typically traveled during each rental contract we relied on some (we believe reasonable) econometric modeling assumptions that we will describe shortly.[2]

The company rents its cars on two types of contracts: a long term contract or a short term contract. Long term contracts are typically written with a maximum duration of one month, combined with a right to automatically renew the previous contract for another month. Rental contracts are at a daily rate with no additional charges for distance traveled during the contract. The daily rate for a long term contract is typically lower than the daily rate for short term contracts. There is a penalty for early returns of vehicles in long term contracts, generally equal to 20% of the lost rental revenue for the unfinished remaining days in the contract.

Fig. 1 illustrates typical rental histories for three different cars in the company's fleet: (1) a compact car rented from one of the company's urban locations, (2) a luxury car rented from an urban location, and (3) recreational vehicle rented from a "tourist" location. In the graphs, a value of '0' denotes a car that is on the lot waiting to be rented, a value of 1 denotes a long term contract, and a value of 2 denotes a short term contract.

We see that the compact and luxury cars that were rented from the urban location started out in a series of long term rentals, with no intervening "lot spells" between the successive monthly rental contracts. It is possible that a succession of unbroken long term contracts might represent a customer who rolled over their monthly contracts into the *de facto* equivalent of a lease, lasting nearly one year in the case of the compact and two years in the case of the luxury vehicle. After these long term contracts came to an end, these vehicles were rented on a series of short term

---

[2] The data include records on the date of accidents and the cost of repairing accident damage, as well as decisions to scrap (versus sell) vehicles that were sufficiently badly damaged as a result of accidents. Although 2543 of the 3908 vehicles in the data set experienced one or more accidents over their service lives, only 123 vehicles were sufficiently badly damaged that they had to be scrapped. In almost all cases where accidents have occurred, the cost of repairing the damage to the vehicle is covered by the insurance of the renter (if the renter was at fault), the insurance of the other party to the accident (if they were at fault), or by the company's insurance (if the party at fault has no insurance). There is a potential indirect source of financial loss due to accidents that the company is not compensated for, namely, if the resale price for cars with accidents is lower. However this effect can be expected to be small, since whenever an accident is repairable, the insurance pays all necessary repairs to restore the car to its pre-accident condition. The company is required to report the number of accidents and information on the nature of each accident (severity, cost of repair and so forth) that a vehicle experienced at the time it is sold. However the econometric evidence we offer below shows that neither the total number of accidents, nor the total cost of repairing these accidents is a significant predictor of resale prices.

Table 1
Regression results for dependent variable IRR

| Variable	Compact All locations		Luxury All locations		RV All locations	
	Estimate	(t-stat)	Estimate	(t-stat)	Estimate	(t-stat)
Constant	0.575	(2.33)	−0.006	(−0.02)	0.999	(1.74)
Utilization rate	0.003	(0.02)	0.522	(4.14)	1.366	(4.52)
Fraction rented long term	−0.220	(−2.52)	−0.076	(−0.88)	−0.876	(−4.79)
Total maintenance costs ($000)	$-7.46e^{-5}$	(−2.81)	$-2.00e^{-5}$	(−1.64)	$6.978e^{-6}$	(0.22)
Odometer (000 km)	0.0007	(1.23)	−0.0004	(−0.56)	−0.001	(−0.79)
Age at sale (years)	0.151	(3.98)	0.072	(1.54)	−0.154	(−1.11)
New price ($000)	−0.104	(−4.98)	−0.036	(−4.22)	−0.082	(−3.22)
Sale price ($000)	0.008	(0.37)	−0.002	(−0.23)	0.063	(1.68)
Short term rental rate	0.003	(2.83)	0.0006	(2.07)	0.004	(3.84)
Long term rental rate	0.037	(19.09)	0.020	(7.25)	0.009	(1.12)
Observations, R^2	167	0.806	40	0.776	31	0.859
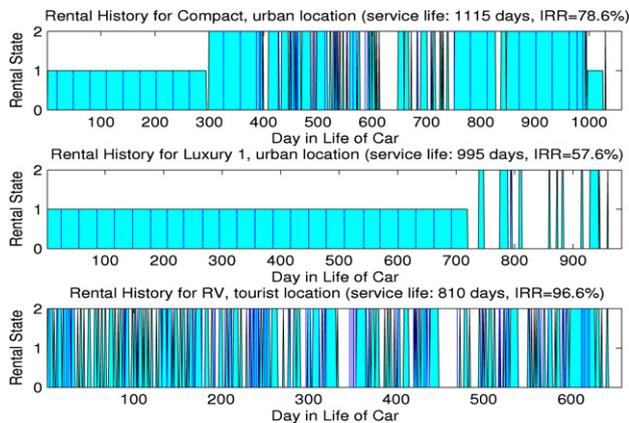
Fig. 1. Typical rental histories for three cars in the company fleet.

contracts, except that the compact car was rented for a final long term contract episode for 30 days near the end of its service life. On the other hand, all of the rentals of the recreational vehicle (RV) in the tourist location were short term rentals, with most contracts lasting only a few days.

Fig. 1 also shows the exact service life and the realized internal rate of return (IRR) that the company earned on the vehicle over its service life. The IRR is defined at the discount rate r (where r is measured on an annual basis) that sets the net present value of the cash flow stream earned by the company over the vehicle's service life equal to zero:

$$0 = \sum_{t=0}^{T} \exp\{-a_t r / 365\} c_t, \tag{1}$$

where T is the number of days over which cash inflows or outflows occurred for the vehicle, c_t is the cash inflow (if positive) or outflow (if negative), and a_t is the number of days after the initial purchase of the vehicle that the tth cash flow occurred. Thus, $c_0 < 0$ and $a_0 = 0$ represent the initial purchase of the car, and then subsequent cash flows would be rental revenues received when the car returned at the end of each rental contract, and cash outflows for maintenance on the dates they occurred. The final cash flow, $c_T > 0$, is the resale price the company receives from selling the car in the used car market, or at an auction. Thus, a_T represents the service life, i.e. the actual age of the car in days at which it was sold, assuming its initial age was $a_0 = 0$ (since all cars were purchased brand new).

We see that for each of the cars illustrated in Fig. 1, the realized rates of return are extraordinarily high. The firm earned a 78.6% rate of return on the compact car, a 57.6% rate of return on the luxury car, and a 96.6% rate of return on the recreational vehicle. The undiscounted profits are also high – $16,683 for the compact,

$24,753 for the luxury car, and $27,654 for the RV – especially in relation to the initial purchase prices of these cars: $9011, $22,808, and $17,889, respectively. The odometer values (in kilometers) on these cars at time of sale were approximately the same, 66,300, 61,000, and 63,265, respectively. However the depreciation rates experienced in the resale values (i.e. the ratio of resale price to new price) of the three cars was quite different: 39%, 56%, and 56%, respectively. The fact that the compact experienced relatively greater price depreciation could be due to having been driven longer (with a terminal odometer of 66,300 and service life of 1115 days, it was approximately 10% older at time of sale than the other two vehicles), or it could just reflect a lower level of durability and thus a greater level of price depreciation for any given odometer value.

The internal rates of return earned by these three example cars in Fig. 1 are not atypical: the mean IRRs for all compacts, luxury, and RV's of the same make, model and vintage as these were 77%, 49%, and 53%, respectively. Table 1 presents the results of a regression of the internal rate of return on various explanatory variables to see which factors are most important predictors of high returns for rental vehicles. We report three regressions for the vehicle types: compact, luxury, and recreational vehicle, pooling over all rental locations.[3] The predicted signs of the coefficients are mostly consistent with intuition: the utilization rate should have a positive coefficient for reasons discussed above, maintenance costs and the new purchase price should have negative coefficients, the sale price should have a positive coefficient, and the daily rental rates for long and short term rental rates should have positive coefficients.

However, we are unable to draw any clear conclusions about the effect of age and odometer value on the IRR: in some cases the coefficients of these variables are positive, and in others negative, and the coefficient estimates are generally statistically insignificant. The results for the maintenance cost variable are also ambiguous. There are a number of possible reasons why the coefficients of age, odometer, and maintenance costs have variable signs and are frequently statistically insignificant. One reason is that these variables have a high degree of collinearity, especially age and odometer. When we re-run the regressions and include only age or odometer individually, the results are still ambiguous, and the coefficients are generally statistically insignificant. Only in one case, for the luxury vehicle, are both age and maintenance statistically significant when odometer is omitted from the regression, and we see in Table 1 that age has

---

[3] The results are basically unchanged if we use the logarithm of the internal rate of return as the dependent variable: the R^2 statistics are slightly lower but the same pattern of signs and significance levels for the coefficients emerges for this alternative specification for the dependent variable in the regression.

a positive coefficient and total maintenance cost has a negative coefficient. But even in this case, the effect of age on IRR is small: the regression results predict that keeping a luxury car for 100 more days increases the IRR by only 0.03 (i.e. by about 5% of the mean IRR of 0.533).

One potential interpretation of the small and statistically insignificant coefficients on age and odometer is that it is an indication of optimizing behavior by the firm. That is, if the firm is choosing age and/or odometer value approximately optimally, we would expect that any variations in these variables about their optimal values should be small. Let $\Pi(o)$ denote the expected discounted profits from keeping a car until it reaches the odometer threshold o before selling it. The optimal odometer threshold o^* is the solution to

$$\frac{\partial \Pi}{\partial o}(o^*) = 0. \tag{2}$$

It follows that if the odometer values at which the company sells its cars are approximately equal to the optimal threshold o^*, we would not detect any significant effect on discounted profits from small variations in the realized odometer value about its optimal value o^* at the time the car is sold. Since IRR is monotonically related to discounted profits, it follows that if the firm is behaving approximately optimally, the effect of small deviations in odometer value from o^* on IRR should also be approximately zero.

However there are a number of reasons why this interpretation may not constitute convincing evidence of optimal behavior on the part of the company. First, as we will show in the next section, the range of odometer values at which the company replaces its vehicles is very wide, more than 100,000 km wide. The argument we made above will only be valid for relatively small deviations of o from its optimal value o^*, and a 50,000 km deviation on either side of o^* seems too large for our argument to apply.[4]

Indeed, as we will show in Section 5, the optimal threshold is generally not a single value o^*, but rather a *function* $o^*(d, r)$ that depends on the rental state of the car r and the duration d in this state. Under the optimal policy (which we calculate by numerical dynamic programming) there is a wide range of odometer thresholds at which replacement can be optimal depending on the values of (d, r). While it is true that small deviations of replacements about $o^*(d, r)$ have small effects on *infinite horizon discounted profits,* it does not follow that such deviations necessarily have small effects on the internal rate of return of the *current car,* since this ignores the effects of delaying replacing the current car on the stream of profits from the car that replaces the current car, the one after that, and so on into the infinite future. Clearly, if we fail to consider that the firm is interested in maximizing expected discounted value of the *stream of current and future profits* and not just the profits on the currently held vehicle, we can get very misleading results. In particular, keeping the current car longer will nearly always increase profits on the margin since even for older cars incremental maintenance costs from keeping a current car a little bit longer are generally far lower than the incremental rental revenues, but keeping the current car longer comes at an opportunity cost in terms of higher profits that might be earned by replacing a currently held old car with a new one. We need to calculate the infinite horizon profit

function (i.e. the value function) to properly account for this trade-off, and this is why a naive approach to calculation of an optimal replacement threshold o^* in Eq. (2) is likely to be misleading.

There is also reason to believe that the coefficient estimates for age and odometer in Table 1 are untrustworthy because these variables are *endogenous.* That is, the company's replacement decisions clearly determine how old and how high the odometer is on its vehicles before they are replaced. If there are unobserved factors associated with a car that lead it to be more profitable (i.e. have higher IRR) these same factors could also lead the company to want to keep the car longer. As a result, one might expect that age and odometer to be positively correlated with unobserved factors affecting profitability and IRR, and this correlation can lead to a spurious upward bias in the coefficient estimates for age and odometer value.

As a result, it is difficult to draw any firm conclusions from Table 1 about whether the company is behaving approximately optimally or not. We would need some sort of *instrumental variable* to deal with this endogeneity problem, but there are no obvious candidates for valid instruments in the data set. What we want would be one or more variables that resulted in *exogenous* shifts in the age at which the company replaced some of its vehicles. An example of such a variable might be a *recall variable,* that is, if there was some major problem in one of the types of cars that the company owned that leads to a recall to the manufacturer, or convinces the company to sell these vehicles before it had intended to sell them. In such case, the "premature" sales of the vehicles could be regarded as a "quasi experiment" that could provide information on how exogenous reductions in vehicle age or odometer values at time of sale would affect the IRR. Unfortunately, there were no recalls of vehicles in the dataset.

The only alternative instrument that we are aware of is an *accident dummy.* If we assume that accidents are purely random events, then an accident is a premature truncation in the intended lifespan of a rental vehicle and serves as a *de facto* randomized experiment that can be used to infer the effect of adopting a shorter replacement threshold on profitability. The average internal rate of return on the 123 vehicles that were scrapped due to accidents in the data set was 36%, which is statistically significantly below the 47% average internal rate of return for all cars in the data set. However there are reasons to be suspicious of accidents as an instrumental variable: if there are "lemons" problems (unobserved problems with certain cars) that make certain vehicles both less profitable and less safe to drive, then the occurrence of an accident could be an indicator of a lemon, and not necessarily evidence that replacing cars earlier reduces profits. Personally, we find this argument a bit far-fetched, but the other major limitation of the instrumental variable approach is something we find more compelling: the accident "experiments" only tell us the effects on profitability of replacing cars too early, *but it does not tell us how profits would change if the company kept its cars longer then it currently does.*

Thus, there are are only two other remaining possibilities for how we might go about testing the hypothesis that this company is a profit maximizer. One is to undertake one or more *controlled experiments,* that is, to pick one or more car types at one or more of the company's locations, and randomly assign some cars to the *treatment group,* where the "treatment" would correspond to keeping cars longer that the company currently does, and cars in the *control group* would continue to be subject to the company's existing replacement policy. By following the cars in the treatment and control group for a sufficient length of time (i.e. from their initial purchase until they are sold), we can compare their profits/returns. If the cars in the treatment group have higher average profits or returns, this would constitute evidence against the hypothesis that the company's existing operating policy is optimal (i.e. profit maximizing).

---

[4] The argument could also be made that the company is choosing an optimal replacement age a^*, but as we will see, there is also a wide range of ages over which the company replaces its vehicles. So the same problem would apply if we hypothesized that the company's replacement threshold was defined in terms of vehicle age rather than odometer value.

The drawback of controlled experiments is that they are costly and time-consuming. Further, there are many possible "treatments" that one could imagine testing: the treatments could involve various combinations of keeping cars longer for various durations in terms of age or odometer relative to the *status quo* as well as various choices about how to discount rental rates for older vehicles. We need sufficiently many vehicles in the treatment and control group to make statistically significant inferences, so the number of possible experiments that the firm could undertake at any point in time is strictly limited. For these reasons, it appears that an experimental approach to testing whether the company is profit-maximizing is not very promising.

The only remaining approach (at least of which we are aware) is to construct an *econometric model* of the firm's operations. This model can be *simulated* to generate predicted outcomes both under the *status quo* and under a variety of alternative hypothetical replacement and operating strategies. The key advantage of the modeling/simulation approach is that simulations are very cheap, and a large number of alternative scenarios and operating strategies can be evaluated extremely rapidly. The key limitation to this approach is that if the econometric model does not provide a good approximation to the actual operations of this company, its predictions of the effects on the firm's profits from implementing various hypothetical alternative operating strategies will not be trustworthy.

In this paper we adopt the modeling/simulation approach. In the next section we present the econometric model of the company's operations, and in the section after that we simulate the model and show that it provides a good approximation to the actual outcomes for this company under its *status quo* operating strategy. Thus, we argue that the modeling/simulation approach is trustworthy, although we still recommend that the predictions of the model be *validated* by conducting a controlled experiment to evaluate whether the optimal replacement strategy implied by this model really does lead to the significant increase in profits that the model predicts.

3. An econometric model

In order to get more insights into the behavior of the rental company and to evaluate the profitability of its vehicle replacement decisions, this section describes an econometric model of the company's vehicle rental operations. We introduce a *semi-Markov model* in which cars that the company owns can be in one of four possible states at any given time: (1) in a long term rental contract (i.e. a "long term rental spell"), (2) in a short term rental contract (i.e. a "short term rental spell"), (3) in the lot waiting to be rented when the previous rental state was a long term rental spell, and (4) in the lot waiting to be rented when previous rental state was a short term rental spell. We refer to the latter two states, 3 and 4, as *lot spells.* We differentiate between these states since it turns out empirically that the duration distribution of a car in a lot spell is quite different depending upon whether it had previously been in a long or short term rental contract.

A *semi-Markov process* is a stochastic process that can be in one of a finite number of possible states at any given time, but where the duration distributions in each of these states (also called the holding time distributions) can be arbitrary distributions. A *Markov process* is a special case of a semi-Markov process where the duration distributions in each state are restricted to be exponential (or geometric, in the case of discrete time models). In this study we use a discrete time model, with the relevant time unit being one day. Let r_t denote the rental state of a given car on day t. From

the discussion above, r_t can assume one of the four possible values $\{1, 2, 3, 4\}$.[5]

In addition to the rental state, other relevant state variables for modeling the decisions of the rental company are the vehicle's *odometer value,* which we denote by o_t, and the *duration in the current rental state,* which we denote by d_t. Thus, we seek to model the joint stochastic process $\{r_t, o_t, d_t\}$. There is another potential state variable of interest, the vehicle's *age* which we denote by a_t. If we let $t = 0$ denote the date at which a car was bought, and if $a_0 = 0$ (when a car is acquired, it is a brand new car), then we have $a_t = t$, i.e. the age of the car in days is the same as the time index t.

In the empirical analysis below, it turns out that a vehicle's age t is strongly correlated with its odometer value o_t. Because of this "collinearity problem" it is difficult to identify the independent effects of these two variables on decisions to sell a car, or on maintenance costs, state transition probabilities, durations in states, and even on the resale price of used vehicles. Since there are numerical and computational advantages to minimizing the number of different variables we include in the dynamic programming model, we have opted to exclude vehicle age from the list of variables that we use to predict the company's selling decision, vehicle resale prices, transitions and durations in spells, and so forth. However, the model does in fact keep track of vehicle age, and as we demonstrate in Section 4, simulations of the model accurately predict the mean age at which the company sells vehicles, even when we restrict the model of the company's sales policy to depend only on the vehicle's odometer o_t and not its age a_t.

Using the three key variables $\{r_t, o_t, d_t\}$ we will also be able to simulate realizations of *rental revenues* and also *maintenance costs* using the data that the company provided us. With this information, we can construct a complete econometric model of the company's rental operations, and conduct stochastic simulations of the model to see how accurately it can represent the company's actual operations. The econometric model requires us to specify and estimate the following objects: (1) a model of the resale price the company receives if it were to sell one of its cars, (2) a model of the random durations of a car in each of the rental and lot states, (3) a model of a car's transition to the next rental spell at the end of the current rental or lot spell, (4) a model of the utilization (kilometers driven) on a particular car during a long or short term rental contract, (5) a model of rental revenues received and maintenance costs incurred by the company over the life of the car, and (6) a model of the company's *selling decision.*

We will now discuss each of these components in turn, describing the econometric model and our empirical findings. The first model is a regression equation to predict the resale price of a rental car when it is sold. We have data on both the new price $\overline{P}(\tau)$ as well as the realized sales price $P_t(o_t, \tau)$ of each car, where τ denotes a particular make and model of vehicle, which we will also call a *car type.* The econometric analysis will focus on the three car types discussed in Section 2, i.e. a particular make/model of compact, luxury, and RV. We wish to emphasize that in order to maintain confidentiality of the data, we are not able to disclose the specific brand and model of these three car types, and instead use the rather vague car type designations to refer to them. Thus, *whenever we refer to one of these car types, such as "compact", we are not referring to the class of all compact cars owned by this company, but instead to a specific brand and model of compact.*

---

[5] Actually we could distinguish a fifth possible state, $r_t = 5$, denoting a brand new car that is in its first lot spell. Empirically we have found that the duration distributions for the initial lot spell can be well approximated as a mixture of the duration distributions for lot spells $r_t = 3$ and $r_t = 4$, and so to reduce the size of the state space, we use only four possible values for r_t and probabilistically assign new cars to lot states 3 and 4 in such a way that the initial duration distribution closely matches the distribution of initial lot spells that we observe in the data.

Table 2
Regression results for dependent variable $\log(P_t(o_t, \tau)/\overline{P}(\tau))$

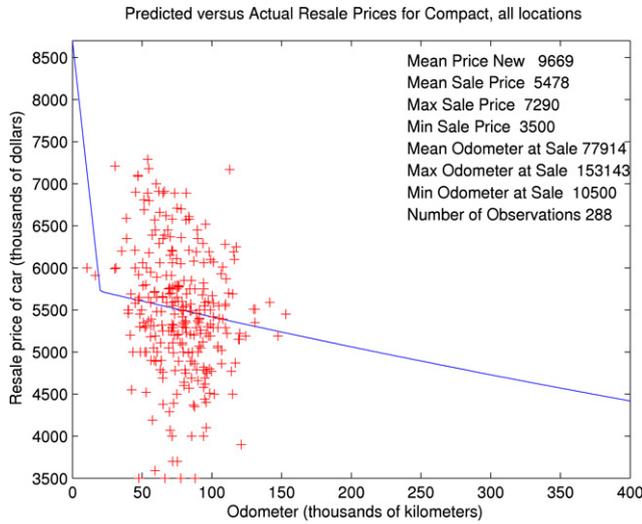| Variable	Compact All locations		Luxury All locations		RV All locations	
	Estimate	(t-stat)	Estimate	(t-stat)	Estimate	(t-stat)
Constant	−0.4789	(−7.61)	−0.6201	(−20.85)	−0.8521	(−4.04)
Age (days)	−0.0001	(−2.53)	−0.0004	(−5.67)	−0.0004	(−2.17)
Odometer (000 km)	−0.0007	(−2.11)	−0.0011	(−2.10)	0.0016	(1.91)
Number of accidents	−0.0112	(−1.05)	0.0006	(0.04)	0.0371	(1.00)
Accident repair costs	−0.8.88e^{-6}	(−1.04)	−4.672e^{-6}	(−0.57)	−1.654e^{-6}	(−0.56)
Internal Rate of Return	0.1629	(12.21)	0.067	(0.99)	0.394	(4.43)
Maintenance cost per day	0.0092	(0.64)	−0.0039	(−0.31)	−0.0053	(−0.33)
N, R^2	288	0.389	91	0.420	41	0.481

Fig. 2. Predicted versus actual resale prices: Compact — all locations.

For each of the three car types τ, we estimated a simple linear regression model with the logarithm of the depreciation rate, $\overline{P}_t(o_t, \tau)/\overline{P}(\tau)$, as the dependent variable

$$\log\left(\frac{P_t(o_t, \tau)}{\overline{P}(\tau)}\right) = \alpha_1(\tau) + \alpha_2(\tau)o_t + \epsilon_t. \quad (3)$$

The results from this model can be interpreted as a regression with type-specific "depreciation coefficients" $(\alpha_1(\tau), \alpha_2(\tau))$ where α_2 measures the effect of the car's odometer value on its resale price. We also estimated regressions where we included the vehicle age and other variables, such as the number of accidents and the total accident repair cost as predictors of the resale price of a car. These results are presented in Table 2.

The constant term in the regressions is a measure of how much depreciation a vehicle experiences the "minute it goes off of the new car lot". We see that this predicted "instantaneous depreciation" is huge for all three vehicle types, but is significantly lower for the compact, which retains 62% of its original value (=exp(−.48)) the instant it is driven off the lot, compared to only 52% for the luxury vehicle and 43% for the RV. Fig. 2 provides a scatter plot of the resale prices for the compacts, graphed against their odometer value at time of sale (the results for the luxury and RV are similar but not shown due to space constraints). The rapid early depreciation in car prices is evident in these graphs. While a number of cars are sold quite "early" after their initial purchase (measured either in terms of their age or odometer value), we do not have any observations of sale prices the company might have received if it were to have sold vehicles in only a matter of a few weeks or months after the initial purchase. For the purposes of our modeling, we did not feel we could trust the regression extrapolations for used vehicle prices for age or

odometer values very close to zero. Therefore, we made a simple linear extrapolation to predict the prices that very new used cars would sell for (i.e. used cars with fewer than 10,000 miles). Fig. 2 show that even for relatively low odometer values the regression accurately predicts the mean resale price. It is immaterial whether we use a straightline extrapolation or assume that the resale function has a discontinuity at zero, since the company almost never sells cars that have fewer than several thousand kilometers on their odometers.

Next we consider the econometric estimation vehicle usage during rental spells. The company does not have mileage charges, and places no constraints on its customers' choice of how far to drive during their rental contracts. The intensity of utilization by rental customers is obviously an important consideration because it determines how quickly a car will "age" in terms of its odometer value, and the odometer is in turn a a key predictor of the resale value of the car. However the difficulty, noted above, is that the firm frequently does not accurately record the in and out odometer values for its vehicles, making it hard for us to determine how far a car was driven on particular rental spells. To get around this problem and make inferences about the conditional probability distribution of the number of kilometers driven of a rental contract of type $r \in 1, 2$ and duration d, we need to impose some additional assumptions.

Let $F(o'|o, d, r)$ denote the conditional distribution of the (frequently unobserved) odometer value on a rental car that has returned from a rental contract of type r, lasting d days, when the out odometer value was o (i.e. the car had an odometer reading of o at the start of the rental spell). Thus, $\nabla o = o' - o$ is the number of kilometers driven by the customer during the rental spell. We assume that the number of kilometers traveled each day by a rental customer are *IID* (Independent and Identically Distributed) draws from an exponential distribution with parameter λ_r. Conditional on spell length d, it follows that $F(o'|d, r)$ is a gamma distribution, since a sum of *IID* exponential random variables has a gamma distribution.[6] The probability density function corresponding to F is given by

$$f(o'|o, d, r) = \begin{cases} \dfrac{[o' - o]^{d-1} \exp\{-(o' - o)/\lambda_r\}}{[\lambda_r]^d \Gamma(d, \lambda_r)} & \text{if } o' - o \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

where $\Gamma(d)$ is the *Gamma function*. Thus, $o' - o$ is the actual number of kilometers traveled by the customer during the rental contract. We have $E\{o' - o|o, d, r\} = d\lambda_r$, so we can interpret λ_r as the mean number of kilometers traveled per day in a rental contract of type r. For notational consistency, we set $\lambda_r = 0$ if $r > 2$, i.e. cars do

---

[6] Actually, the distribution is part of a special subclass of the Gamma family known in renewal theory as the *Erlang distribution* since the parameter α of the Gamma distribution is an integer $\alpha = d$.
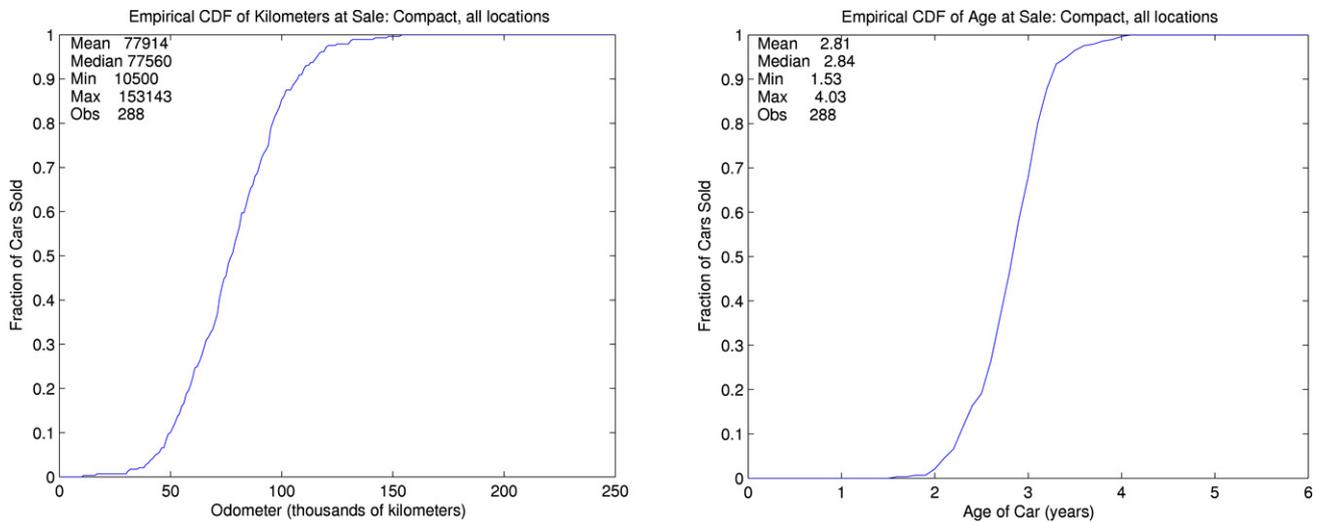
Fig. 3. Cumulative fractions of cars replaced: Compact — all locations.

not travel any kilometers when they are on the lot waiting to be rented.

In order to estimate kilometers traveled per day under short term and long term contracts, it would be natural to look to the rental contract data directly and take the average kilometers traveled per day for short term and long term contracts separately. Since the in and out odometer values for rental contracts do not appear to be accurately recorded in the company's data, we cannot use this approach. If we had such data, it might even be possible for us to estimate the conditional distributions $F(o'|o, d, r)$ non-parametrically or semi-parametrically, and thus, not have to rely on the parametric assumption that kilometers traveled per each day of a rental contract are *IID* exponential variables. In the absence of contract-specific mileage data, we can still estimate the two λ_r parameters necessary for us to determine the distributions $F(o'|o, d, r)$ using the accurate records we do have on the odometer value of each vehicle at time of sale.

Suppose that at time of sale, a rental car had been rented for N^s days under short term rental contracts and N^l days under long term rental contracts. Then the odometer value on the car is given by

$$\tilde{o} = \sum_{i=1}^{N^l} \nabla o_i^l + \sum_{i=1}^{N^s} \nabla o_i^s \tag{5}$$

where ∇o_i^l and ∇o_i^s are the realized number of kilometers traveled under long and short term contracts, respectively. Under our assumptions that kilometers traveled per day are exponential random variables with parameters λ_1 (for long term contracts) and λ_2 (for short term contracts), we have

$$E\{\tilde{o}|N^l, N^s\} = \lambda_1 N^l + \lambda_2 N^s. \tag{6}$$

Since we do observe the odometer at sale $\tilde{o}$ and number of days a vehicle is rented, N^l and N^s, this implies that we can estimate λ_1 and λ_2 as coefficients on a simple linear regression

$$o_i = \lambda_1 N_i^l + \lambda_2 N_i^s + \varepsilon_i \tag{7}$$

where o_i is the odometer at time of sale on the *i*th rental car sold by the company, and N_i^s and N_i^l are the number of days the *i*th car had been in short and long term rentals over its service life. These results are presented in Table 3.

We use the regression estimates and the information on (N_t^l, N_t^s) for each car, now indexed by the day in its service life, t, to compute a predicted value for the car's odometer, $\hat{o}_t =$

Table 3
OLS estimates of λ_1 and λ_2 in the regression equation (7)

Variable	Compact	Luxury	RV
	All locations	All locations	All locations
λ_1	78.7	86.6	95.4
λ_2	157.1	140.8	167.7

$\hat{\lambda}_1 N_t^l + \hat{\lambda}_2 N_t^s$, at day t in the car's life. The high R^2 values for the odometer regressions in Eq. (7) give us confidence that the predicted odometer values are reasonably accurate.[7] With $\hat{o}_t$, the next step is to analyze the determinants of the company's decision to sell its cars.

We estimated a reduce-form binary logit model to capture the company's *status quo* replacement policy. We do not report the detailed results here: it suffices to note that the only significant predictors are the age and odometer of the vehicle. Fig. 3 summarizes the firm's replacement policy for compact cars (the replacement policy for luxury and compact is very similar). It shows the cumulative distribution for replacements as a function of the odometer value (left hand panel), and the cumulative distribution of the vehicle age (right hand panel). The right hand column confirms the company's claim that the target replacement age for its vehicles is 3 years. The mean age of the three types of cars at replacement is fairly close to this three year target: 2.8, 2.9 and 2.7 years for the compact, luxury and RV, respectively. However the left hand panel shows that in terms of odometer values at replacement, there is greater variability. The mean odometer value at replacement for the three vehicle types is 78, 75 and 89 thousand kilometers, respectively. The fact that mean replacement ages vary much less across the three car types than the mean odometer values at replacement may be taken as evidence that the company bases its replacement decision more on the "three year rule" than on a rule based on number of kilometers driven. As we will see in the next section, due to the high degree of collinearity between age and odometer values, a replacement rule based on odometer value can provide a good approximation to an age-based replacement rule and vice versa.

---

[7] We also tested the accuracy of the regression predictions by comparing the actual odometer reading that was recorded at dates where maintenance was performed with the predicted values $\hat{o}$ on these same dates. The distribution of prediction errors is centered at zero with small variance, indicating our regression model is a very good predictor of actual vehicle usage.

The remaining objects to be estimated are the *spell durations* and the *spell transition probabilities.* As is well known, there is a duality between duration distributions and the corresponding *hazard functions.* We chose to work with hazard functions and let $h(d, r)$ denote the hazard rate for the rental state r, i.e. it is the conditional probability that the car that has been in rental state r for d days will exit the state d on the next day, $d + 1$, and with the complementary probability $1 - h(d, r)$ it will continue to remain in state r. Since we have sufficiently many observations of rental spells, we were able to estimate the hazard functions for these spells non-parametrically. The longest duration for any rental spell is 31 days, i.e. the maximum duration of a monthly rental. As one might expect, the duration distributions of long and short term rental spells are very different: most short term rentals last only a few days, whereas most long term rentals last for an entire month. There is only minor variation in the durations of long term rentals, i.e. some rentals are for 29 days, 30, or 31 days. Most long term rentals last more than 15 days, perhaps in part due to the 20% penalty the company imposes on early return of vehicles in a long term contract.

We have far fewer observations on lot spell durations, especially for type 3 lot spells, i.e. where the previous rental spell was a long term contract. This is due to the high probability of roll overs in longer term contracts, leading to relatively few observations on intervening lot spells with positive durations. Due to the relatively small number of observations, the non-parametrically estimated hazard functions are quite jagged. Also, unlike rental contracts, there is no *a priori* upper bound on the duration of a lot spell. As a result we needed some method of extrapolation to predict durations given that we have only a small number of cases with extremely long lot durations. Our solution to this problem was to assume that the hazard function is constant after $d = 31$ days, which implies a geometric upper tail for the distribution of lot spells. We estimated this constant upper tails of the hazard function by imposing the constraint that the implied duration distribution (with a smoothed, non-parametrically estimated lower tail and the geometric upper tail) has a mean duration that equals the actual mean duration for type 3 and 4 lot spells, respectively.

When a spell in a given rental state ends, there is a transition to a new rental state. Let $\pi(r'|r, d, o)$ denote the probability that the new rental state for a car will be r' given that the current rental state is r, the odometer value is o, and the duration in state r is d. We call π the *rental state transition probability.* Recall that if $r > 2$, the car is in a lot spell. We rule out the possibility of "self transitions" to the lot (that is, we assume that $\pi(r|r, d, 0) = 0$ for $r > 2$) because the hazard function $h(d, r)$ already provides the probability that the lot spell has ended, and there is no conceptual difference between a lot spell continuing for one more day, versus the case where a lot spell terminates and immediately re-enters the lot via a self-transition $r' = r$. Thus the restriction $\pi(r|r, d, o) = 0$ for $r > 2$ can be viewed as an econometric identification normalization.

However, for rental spells, there is a conceptual distinction between a rental spell that terminates with an immediate transition to a new rental spell versus the case where an existing rental contract continues for one more day. The former case can be viewed as an immediate roll over of one rental contract to another one, such as when a previous customer renews or extends their previous rental contract by one more month (in the case of a long term contract), or by another day (in the case of a short term contract). Thus, we allow $\pi(r|r, d, o) > 0$ for $r \in \{1, 2\}$, and interpret this probability as a probability of a contract extension or roll over.

The rental spell transition probability can also accommodate transitions from a rental spell to a lot spell, except that by our

definition of the two types of lot spells, it must be the case that $\pi(4|1, o, d) = 0$ and $\pi(3|2, o, d) = 0$, i.e. if a car is leaving a long term rental spell, it can only transition into a lot spell of type 3 which is defined as a lot spell where the previous rental spell was a long term contract. Similarly, a car leaving a short term rental spell can only transition into a lot spell of type 4. The reason why we distinguish the two types of lot spells is that the hazard functions and mean durations for type 3 lot spells are different than for type 4 lot spells. In particular, for all three types of cars, hazard rates for type 3 lot spells are lower and mean durations are higher. In plain language, if a car had previously been in a long term rental that did not immediately roll over, one would expect the car to be on the lot for a longer period of time compared to the case where the car has returned to the lot from a previous short term rental.

Since there are three possible destination states for transitions out of a rental spell i.e. immediate return to a long term or short term rental, or back to the lot, we used a trinomial logit model to estimate these probabilities. This probability is given by

$$\pi(r'|r, d, o) = \frac{\exp\{v(r, d, o)\theta_{r'}\}}{\sum_{\rho \in \{1, 2, l(r)\}} \exp\{v(r, d, o)\theta_\rho\}}, \tag{8}$$

where $v(r, d, o)$ is a vector-valued function of the variables (r, d, o) and θ_ρ is an alternative-specific vector of parameters, for $\rho = \{1, 2, l(r)\}$ (where $l(r)$ denotes a lot spell, either of type 3 if $r = 1$ or type 4 if $r = 2$) with the same dimension as v. As is well known, it is not possible to identify all three of the θ_ρ vectors. Therefore we make an identifying normalization that $\theta_1 = 0$, i.e. we normalized the parameters for transition to long term contract to zero.

Since our identification normalization rules out "self-transitions," there are only two possible destinations from a lot spell: i.e. to either a long term or short term rental spell. We used a binary logit model to estimate transition probabilities out of lot spells. Similar to the trinomial logit specification, we used transition probabilities specified as

$$\pi(r' = 1|r, d, o) = \frac{\exp\{v(o, d)\theta_r\}}{1 + \exp\{v(o, d)\theta_r\}}, \quad r \in \{3, 4\}. \tag{9}$$

We do not report the actual estimation results here, but there are two main findings we think are noteworthy: (1) for all car types, there is a very high probability that cars will be initially rented in long term contracts, (2) the results provide clear evidence of "contract age effects". That is, as the odometer value increases (i.e. the age of the car increases) the probability of transitions into long term rental contracts decreases, the probability of transitions into short term rental contracts increases, and the probability of remaining on the lot unrented increases.

The remaining objects in the econometric model of this company are the daily rental rates and maintenance costs. Daily rental rates do not need to be estimated since we can simply access the company's published tariff rate. However there is a small amount of variability in daily rental rates due to variations in optional equipment and features on cars (e.g. some cars have larger engines that the standard size, etc.). To account for this variation, we simply computed the mean daily rental rate by dividing the total rental revenues earned in long and short term rental spells by the number of days in these spells.

Maintenance costs are incurred on a episodic basis. The company appears to adhere to a fairly regular periodic maintenance schedule with operations such as oil changes, brake pad replacement and so forth occurring at regular intervals (either intervals of time such as every 3 months, or in terms of odometer values, such as every 50,000 km, etc). However, there is also evidence of a high frequency of "unexpected maintenance" resulting from random malfunctions or problems in particular cars. While we could have tried to model the durations between successive maintenance

Fig. 4. Simulated versus actual odometer and age at sale: Luxury — all locations.

visits, with a conditional distribution of maintenance costs that are incurred at each maintenance event, we opted for a simpler approach that appears to work just as well. Our approach is simply to charge a daily equivalent maintenance rate, where we estimate the daily maintenance charge by taking the mean of the ratios of total maintenance costs over the service life of the vehicle divided by the age of the vehicle at the time it was sold. In the next section we will now show, via the stochastic simulations, that this simplified treatment of maintenance costs does not compromise our ability to provide a good overall model of the firm's operations.

The final issue we address is whether there is evidence of aging effects in vehicle maintenance costs. Scatter plots of the average daily maintenance costs incurred by the company on its vehicles as a function of the predicted odometer value show no evidence that daily maintenance costs increase with the odometer value of the vehicle, at least over the range of odometer values observed in this company's fleet. Thus, we conclude that the only aging effects that we can detect in the econometric analysis are: (1) the rapid decline in resale values of vehicles as a function of their age and odometer value, and (2) the "rental composition aging effect", i.e. the tendency for cars to be initially rented on long term contracts, but to gradually transition to an increasing share of short term rental contracts and to spend more time on the lot as the vehicle ages.

4. Evaluating the econometric model: Simulated versus actual outcomes

In the previous section we described and estimated an econometric model of the rental company's operations. In order to determine if this is a good model that accurately captures the key features of the behavior of this company, this section presents comparisons of simulated outcomes from the econometric model to the actual outcomes for each of the tree vehicle types analyzed in Section 3.

Our approach to simulating the econometric model is conceptually straightforward. Starting with a new car on the lot with an odometer value of zero, we simulate arrival of customers using the estimated transition probabilities discussed in Section 3. Once a car is rented (either in a long or short term rental spell), we use the estimated hazard functions to determine the duration of the rental spell, and the estimated regression coefficients (from the regression Eq. (7) that determines average daily kilometers traveled in short and long term rental contracts, respectively) to predict

the number of kilometers driven during the rental spell. More precisely, we use the fact that under the assumption that daily kilometers traveled are *IID* draws from an exponential distribution with parameter λ_1 (for short term rentals) or λ_2 (for long term rentals), and the total duration of the rental spell to generate a random draw of the total number of kilometers driven during that spell. We continue this process until the first time a simulated car re-enters the lot with an odometer reading o_t that exceeds its *replacement threshold* $\bar{o}$. If $o_t > \bar{o}$, the car is sold (where we use a draw from a lognormal distribution determined by the used car sale price regression to predict its sales proceeds), a new car is purchased for a price $\bar{P}(\tau)$, and the life of the new simulated car begins. Finally, we draw the replacement thresholds $\bar{o}$ at the start of the life of each newly purchased car, using the empirical distribution of odometer values at which the company actually replaces its cars in the left hand panel of Fig. 3.

Fig. 4 presents comparisons between simulated and actual distributions of the odometer and vehicle age (in days) at which vehicles are replaced. We present the results for the luxury car type only, although the results for the compact and RV are similar. The left hand panel of Fig. 4 compares the actual distribution of odometer values (solid blue line) with the simulated distribution (dashed red line). We see that the two distributions are close to each other, which is a result we would expect given that we have drawn the odometer thresholds that determine when vehicles are to be replaced in the simulations from the actual (empirical) distribution. Thus, the differences in the two distributions in the left hand panel of Fig. 4 are entirely due to *sampling error* in our random sample of 100 simulated cars.

The right hand panel of Fig. 4 compares the actual distribution of replacement ages to the one implied by the econometric model. In this case we do not directly draw the age at which a car is replaced from the empirical distribution of replacement ages, so there is no guarantee that the simulated distribution of ages at replacement is close to the actual distribution. Indeed, the simulated age at replacement is a result of a more complicated set of interactions that depend on other estimated objects in the econometric model that determine the number of times a vehicle was rented, the durations of these rental spells, and the numbers of kilometers driven per rental spell. This implies a particular co-evolution of vehicle age and odometer values, so that when the simulated odometer value exceeds the replacement threshold $\bar{o}$ that we randomly drew from the empirical distribution in the left panel of Fig. 4 at the start of each car's simulated history, the random time at which the car's simulated odometer exceeds $\bar{o}$ determines the simulated lifetime (age) of the vehicle in question.