Inferring the Economic Preference of a Rental Vehicle Company by Modeling Its De-fleeting Process [*,†]

Chuan-Hsiang Han[‡], Jingren Shi[§] and Suzhou Huang[¶]

December 29, 2015

Abstract

When a vehicle manufacturer designs a contract with a rental vehicle company it is important for the OEM to properly understand the rental company's economic preference. While it is usually not directly observable, the economic preference of the counter party can often be revealed indirectly through some observable market behavior. In such cases, econometric inference needs to be used. In this paper, we use the de-fleeting process of the rental vehicle company as the inferential apparatus. To this end, we first develop a model to describe the decision-making in the de-fleeting process for the rental vehicle company, based on the optimal stopping theory. We then outline an econometric procedure to estimate the model parameters. Finally, we use simulated data to illustrate how to deal with some of the technical issues that one might encounter when the procedure is applied to real data.

[*]Authors acknowledge helpful comments and suggestions from the editor-in-chief Professor Michael McAleer.

[†]Opinions expressed herein are entirely those of the authors, rather than those of Ford Motor Company.

[‡]Department of Quantitative Finance, National Tsing Hua University, Hsinchu, Taiwan, 30013, ROC, chhan@mx.nthu.edu.tw.

[§]Research and Advanced Engineering, Ford Motor Company, Dearborn, MI 48121-2053, USA

[¶]Research and Advanced Engineering, Ford Motor Company, Dearborn, MI 48121-2053, USA

1 Introduction

When designing a business contract one of the most important tasks is to assess the economic preferences of the companies involved. This assessment is not a non-trivial exercise when the economic preference is not directly observable. In such an environment, econometric inference techniques need to be applied to certain observable data that can indirectly reveal the relevant information. In this paper we consider an example in which the economic preference of a rental vehicle company is extracted from the data of when the rental vehicle company disposes individual used-vehicle after its fleet service.

From the viewpoint of consumption capacity in tourism, planning for public transportation such as airplanes and car rental are essential. The air transportation has earned enormous studies while feet planning in car rental is very limited [3] despite that car rental is much more liquid with much larger data set than air planes. To our best knowledge, this article is the first paper to analyze the de-fleeting process of a rental vehicle company.

Specifically, we first parameterize the utility-flow function of the rental company for a given vehicle. We then derive a micro-level model that describes when the vehicle should be taken out of the rental service and auctioned in a used-vehicle wholesale market. The values of the parameters are then determined by inverting a micro-level model under standard maximum likelihood framework.

While the design of the contract between a vehicle manufacturer and the rental vehicle company is the ultimate goal, the scope of the paper is limited to the econometric inference task alone. However, issues related to the identifiability of the utility-flow function will be addressed carefully. For business sensitivity reasons and proprietary nature of the data , we will not use the real data in this paper. Instead, we use simulated data to illustrate how the inference is done.

Before proceeding further, it is appropriate to detail the relevant business environment in question. A rental vehicle company buys a fleet of vehicles from a vehicle manufacturer.[1]

[1]We consider only the so-called risk units, i.e., the units that are actually bought by the rental company with property rights transferred. There is another form of transaction that will not be considered here but

These vehicles are then put into consumer-rental business for a period of time. Depending on usage and the current market condition, the rental company decides when to retire a specific vehicle and sell it on the wholesale auction market. At any point in time, the rental company needs to strike a balance between the accumulated utility-flows of the vehicle being rented to consumers and how much the vehicle can be sold for at a secondary market. Since the residual value of the vehicle is stochastic, the decision process of the rental company can be modeled as a process of exercising a perpetual American option with both utility flow and terminal condition. Mathematically, this amounts to solving a classical optimal stopping problem.

One of the underlying assumption of this paper is that the de-fleeting process for each individual vehicle can be treated independently. This is an oversimplification to be sure, but there are supporting arguments that can be made to partially justify the assumption. First, the overall demand level in the industry is largely predictable, so fleet size can be well planned in advance. Second, the rental industry is sufficiently competitive. Any temporary shortfall in total fleet size can be made up quickly without incurring too much extra costs.

Of course, the de-fleeting process depends on many other factors in the rental vehicle business, such as seasonality and occasional industry shocks. These factors will not be explicitly modeled in this paper. Therefore, it is extremely important to remove the effects of factors from the data before applying the theory developed here to an individual vehicle's de-fleeting process.

The remainder of this paper is organized as follows. In Section 2 we present the rental company's decision model. Section 3 addresses the extent to which the model is identifiable econometrically with available data. Due to the proprietary nature of the real data[2] and business sensitivity, we only illustrate how to estimate model parameters using simulated data (Section 4). In the same section, sensitivity analysis is included to illustrate the

is more popular in the industry: repurchase program. Under a repurchase program, a vehicle bought by the rental company is guaranteed to be purchased back by the manufacturer with a specific contract term.

[2]Ford's internal new-vehicle sales data are matched with NADA's used-vehicle auction data at the level of Vehicle Identification Number.

redundancy of the model in parameterizing the utility flow function. We summarize and conclude in Section 5.

2 Rental Company's Decision Model

The residual (value) process of a used vehicle is denoted by $\{R_t; t \geq 0\}$. We model the residual process as a continuous-time geometric Brownian motion

$$dR_t = -\mu\, R_t\, dt + \sigma\, R_t\, dW_t\,. \tag{1}$$

where W_t is the standard Wiener process, and $\mu > 0$ indicating the value of a vehicle generally depreciates over time on average. It is important to emphasize that R_t should be understood as the subjective perception of the residual value of a rental vehicle at time t by the manager who makes the decision to retire the vehicle from its rental service and sell it at wholesale auction. The underlying basis for R_t is related to the stochastic nature of the usage of the vehicle and the pertinent market conditions. The discrepancy between the subjective perception and the realized value in the market will be interpreted as the regression error.

Intuitively, we imagine that the manager has a mental model that attempts to balance the accumulative revenue from renting out the vehicle to consumers against the value he can get if he decides to sell the vehicle at a secondary market. This expectation can be formalized as

$$V(t, R_t = r) = \sup_{\tau \in [t,\infty)} E\left\{\int_t^\tau e^{-\rho\,(s-t)} u(s, R_s; \Theta_1)\, ds + e^{-\rho\,(\tau-t)} h(\tau, R_\tau; \Theta_2)\right\}, \tag{2}$$

where $V(t, R_t = r)$ is the net present value associated with the vehicle; $u(t, R_t; \Theta_1)$ denotes the utility flow (revenue less operating cost per unit of time) when the vehicle is being rented; $h(t, R_t; \Theta_2)$ denotes the terminal value when the vehicle is sold at an auction, ρ is the discount rate of the rental company, Θ_1 and Θ_2 are the parameters that characterize the utility flow function and the terminal condition, respectively. The expectation E in the above equation is with respect to the ensemble of the residual process in Eq.(1). Since the optimal stopping time τ is random, which depends on each realized residual trajectory, the maximization should be understood as adaptive with respect to the residual process R_t, i.e., $\tau = \inf\{t \geq 0; R_t = \bar{R}\}$ is the first hitting time of an exercising threshold $\bar{R}$ to be defined and solved later.

The problem defined in Eq.(2) can be handled by several standard methods. Here we will explicitly follow the dynamic programming approach well described in Dixit and Pindyck (1993). Let Δ denote an infinitesimal increment in time. Imagine that the problem in Eq.(2) is discretized. At any time t with a residual level $R_t = r$ the rental company makes a decision on whether to continue renting the vehicle to consumers over the next period or to sell the vehicle at an auction to gain a terminal payoff $h(t, r; \Theta_2)$. In the case that the decision is continuation, the rental company accrues a utility flow $u(t, r; \Theta_1)\,\Delta$. In addition, the rental company has the opportunity to decide what do do in the next period. Therefore the value function $V(t, r)$ obeys the following Bellman equation

$$V(t, r) = \max\left\{u(t, r; \Theta_1)\Delta + e^{-\rho\Delta}E\left[V(t+\Delta, R_{t+\Delta})\right],\ h(t, r; \Theta_2)\right\}.$$

In the *continuation* region the above equation simplifies to,

$$V(t, r) = u(t, r; \Theta_1)\Delta + e^{-\rho\Delta}E\left[V(t+\Delta, R_{t+\Delta})\right].$$

Applying Ito's lemma and keeping only the leading order terms in Δ yields the following partial differential equation for the net present value function

$$\frac{\partial V}{\partial t} - \mu r\frac{\partial V}{\partial r} + \frac{1}{2}\sigma^2 r^2\frac{\partial^2 V}{\partial r^2} - \rho V + u(t, r; \Theta_1) = 0. \qquad (3)$$

On the boundary between the *continuation* region and the *exercising* region, we have

$$V(t, \bar{R}) = h(t, \bar{R}; \Theta_2).$$

However, the location of the boundary between the *continuation* region and *exerciseing* region is still unknown. That is, we are dealing with a free-boundary problem. To completely specify the solution additional restrictions need to be added. These additional restrictions depend on the detail of the problem at hand (see Dixit and Pindyck (1993)). In our case, one restriction is the so-called smooth pasting condition: $\frac{\partial V}{\partial r} = \frac{\partial h}{\partial r}$ when $r = \bar{R}$. The other restriction is to prevent the speculative bubble: V/r is bounded at $r \to \infty$.

To proceed further, we need to take concrete forms for u and h. We choose the utility flow function to have the simple form: $u(r;\Theta) = \alpha(r/R)^{\nu_1} - \beta(r/R)^{-\nu_2}$, with $\Theta \equiv \{\alpha, \beta > 0;\ \nu_1, \nu_2 > 0\}$, and R is the value when the vehicle is new. The α term is the revenue income per unit, and the exponent ν_1 is to capture the effect that a new vehicle is more attractive to consumers. The β term is interpreted as the unit operating cost, and ν_2 is to account for the fact that maintenance cost increases as the vehicle ages. We choose the terminal condition to have the form $h(r;\kappa) = (r-\kappa)^+ \equiv \max(r-\kappa, 0)$, where κ denotes a fixed transaction cost for selling a used vehicle in the wholesale auction market. Under these choices the complete solution to Eq.(2) is given by the following free boundary problem

$$\begin{cases} -\mu r V' + \frac{1}{2}\sigma^2 r^2 V'' - \rho V + \alpha(r/R)^{\nu_1} - \beta(r/R)^{-\nu_2} = 0, & \bar{R} \le r, \\ V(r) = (r-\kappa)^+, & 0 \le r < \bar{R}, \\ V'(r)\Big|_{r=\bar{R}} = 1, & \\ \Big\{V(r)/r\Big\}_{r\to\infty} = \mathcal{O}(1). & \end{cases} \tag{4}$$

In the above equation we have dropped the time derivative term from Eq.(3) because we are obviously dealing with a time-homogenous problem, and hence are only interested in time independent solution.

The explicit solution to Eq.(4) can be derived using the standard technique. Without going through the detail we give the final expressions here. The net present value function has the form

$$\begin{aligned} V(r) &= A_1\Big(r/R\Big)^{\nu_1} + A_2\Big(r/R\Big)^{-\nu_2} + B\Big(r/R\Big)^{q}, \quad r > \bar{R} \\ V(r) &= (r-\kappa)^+, \qquad \bar{R} \ge r > 0 \end{aligned} \tag{5}$$

where the exponent q and coefficients A_1, A_2 and B are given respectively by

$$q = -2\rho\left[\mu + \sigma^2/2 + \sqrt{(\mu+\sigma^2/2)^2 + 2\sigma^2\rho}\right]^{-1}, \tag{6}$$

$$A_1 = \alpha\left[\rho + \nu_1(\mu + \sigma^2/2) - \nu_1^2\sigma^2/2\right]^{-1}, \tag{7}$$

$$
\begin{aligned}
A_2 &= -\beta\left[\rho - \nu_2(\mu + \sigma^2/2) - \nu_2^2\sigma^2/2\right]^{-1}, &(8)\\
B &= \left[\frac{\bar{R}}{q} - A_1\frac{\nu_1}{q}\left(\frac{\bar{R}}{R}\right)^{\nu_1} + A_2\frac{\nu_2}{q}\left(\frac{\bar{R}}{R}\right)^{-\nu_2}\right]\left(\frac{R}{\bar{R}}\right)^{q}. &(9)
\end{aligned}
$$

One can recognize that q is one of the characteristic roots of the quadratic form associated with the homogenous part of Eq.(4). The homogenous term associated with the other root, which diverges when $\sigma \to 0$, is excluded by the restriction of no speculation bubble. The exercising threshold $\bar{R}$ is solved from the equation

$$
A_1(q - \nu_1)\left(\frac{\bar{R}}{R}\right)^{\nu_1} + A_2(q + \nu_2)\left(\frac{\bar{R}}{R}\right)^{-\nu_2} = (q-1)\bar{R} - \kappa q\,. \tag{10}
$$

Eq.(10) is generally not amenable in closed forms, but can be easily handled numerically. However, for special values of ν, explicit solutions can be obtained. For example, the option exercising thresholds for $\nu_1 = 1$ and $\nu_2 = 0$, or $\nu_1 = 0$ and $\nu_2 = 1$, can be easily derived,

$$
\bar{R} = \frac{\beta - \rho\kappa}{\alpha/R - \mu - \rho}\,\frac{(\mu+\rho)q}{\rho(q-1)}\,, \qquad \text{when } \nu_1 = 1 \text{ and } \nu_2 = 0\,; \tag{11}
$$

$$
\bar{R} = \frac{\alpha + \rho\kappa}{\beta/R + \mu + \rho}\,\frac{(\mu+\rho)q}{\rho(q-1)}\,, \qquad \text{when } \nu_1 = 0 \text{ and } \nu_2 = 1\,. \tag{12}
$$

3 Model Estimation Strategy

From the nonlinear algebraic equation (10), the optimal stopping rule $\bar{R}$ depends on two types of parameters: those specifying the residual process: $X = (\mu, \sigma, R)$, and those specifying the decision-making process $\Theta = (\alpha, \beta, \nu_1, \nu_2, \rho, \kappa)$. Our goal in this section is to develop procedures to estimate these parameters.

3.1 Rental Pricing Peculiarity

A peculiar feature of pricing in the consumer rental vehicle industry is its insensitivity to vehicle details. Prices are typically charged at the vehicle nameplate level, independent of the vehicle series, option content, etc. This kind of simplified pricing convention is true to both situations, when the rental company buys vehicles from OEMs and when the the rental company charges its consumers.

On the other hand, the residual process is highly dependent on vehicle series, option content, seasonality, usage pattern (such as miles driven, wear and tear, and so on). These features can serve as potential covariates in the model parameter estimation process.

3.2 Two-Step Procedure

Since we have two sets of parameters (residual process related and decision-making related) we need a separate procedure for each of them. The general strategy here is to divide all de-fleeted vehicles into groups. The rule of division is such that vehicles in the same group are regarded as obeying the same residual processes, and vehicles in different groups obey different residual process. Then, the residual related parameters for group g, $X^g = (\mu^g, \sigma^g, R^g)$, are estimated using intra-group data; and the decision-making parameters, $\Theta = (\alpha, \beta, \nu_1, \nu_2, \rho, \kappa)$, which are assumed to be shared by all groups (see the aforementioned note on rental pricing peculiarity), are estimated using inter-group data with X^g as covariates. It is easily anticipated that how to properly divide de-fleeted vehicles into groups is critical in analyzing the real data.

3.2.1 Intra-Group Estimation

For each vehicle Group g, we perform the following regression, consistent with the lognormal assumption in Eq.(1),

$$\ln p_i = \ln R^g - \mu_1^g \tau_i + \epsilon_i \,, \qquad \epsilon_i \sim N(0, \tau_i \sigma_g^2) \tag{13}$$

where i labels an individual vehicle in Group g, τ_i is the month-in-service of the vehicle, and $\mu_1^g = \mu^g - \sigma_g^2/2$. This regression is not a standard OLS, but rather a weighted OLS (with weighting factor $1/\tau_i$), due to the τ_i dependence in the error term. This regression yields simultaneously the estimates for all the residual related parameters in $X^g = (\mu^g, \sigma_g, R^g)$. Here R^g should not be confused with the average transaction price of the vehicles in this group between the OEM and the rental company. Rather, it should be understood as the average price had the rental company decided to sell the vehicles in an auction market immediately after acquiring them from the OEM.

3.2.2 Inter-Group Estimation

At the inter-group level, we first aggregate all the individually observed option exercising thresholds within a group according to

$$\bar{R}^g_{\text{observed}} = \frac{1}{N_g} \sum_{i \in g}^{N_g} \bar{R}_i \,, \tag{14}$$

where N_g denotes the number of observed vehicles in the group. Then we perform the following nonlinear least square estimation of $\Theta = (\alpha, \beta, \nu_1, \nu_2, \rho, \kappa)$, using aggregated R^g_{observed} across all groups,

$$\bar{R}^g_{\text{observed}} = \bar{R}_{\text{theory}}(X^g, \Theta) + \epsilon^g \,, \qquad \epsilon^g \sim N(0, \sigma^2_{\bar{R}}) \,. \tag{15}$$

The estimator is defined to minimize the following objective function

$$\hat{\Theta} \equiv \arg\min_{\Theta} \; \frac{1}{G} \sum_{g=1}^{G} \Big(\bar{R}^g_{\text{observed}} - \bar{R}^g_{\text{theory}}(\Theta) \Big)^2 . \tag{16}$$

With the assumption on the error term in Eq.(15), the above estimation is equivalent to maximum likelihood estimation.

3.3 Model Parameter Identification

While the intra-group estimation is straightforward, the inter-group estimation can suffer from severe co-linearity among the parameters in Θ. The origin of the problem can be traced back to the fact that observed data is not very sensitive to some combinations of these parameters in the decision-making process with the intuitive functional form adopted for $u(r) = \alpha(r/R)^{\nu_1} - \beta(r/R)^{-\nu_2}$. This in turn means that we cannot estimate all parameters in Θ using option exercising thresholds alone, unless we have additional information. One plausible strategy to proceed is to first assume that parameters related to costs, such as β

and κ, are estimated from other independent sources, and then check the sensitivity on the assumed values of these cost parameters. If severe co-linearity arises from the estimation process, we should not find a high degree of sensitivity on the assumption for the cost parameters, since co-linearity implies certain redundancy of the parameterization in the decision-making process. As we will see, these expectations will be explicitly borne out in the next section.

4 Numerical Simulations

Given the multi-parameter complex structure of the stopping rule $\bar{R}$ solved from Eqn. (10), it is convenient to reveal certain properties of $\bar{R}$ by means of numerical simulations while analytic approach may not exist. To check the robustness of nonlinear least square estimation, we disturb a given set of model parameters, which are comparable to real data consensus, by a white noise, then examine those estimated model parameters with the actual ones and possible identification problems.

In this section we illustrate all the relevant issues in the estimation process using simulation. The other purpose here is to amass practical experience in a controlled environment before we tackle real data, which are inevitably noisier than simulated data. Since the intra-group level estimation is trivial, we will only concentrate on the inter-group estimation. For simplicity, we further limit to the case of $\nu_1 = \nu_2 = \nu$ in this section.

4.1 Simulated Data

We choose the decision-making parameters to have the value of $\Theta = (\alpha = 1800, \beta = 1000, \nu = 0.4, \rho = 0.015, \kappa = 300)$. These numbers are motivated by the real data in terms of their order of magnitude. We then generate $N = 100$ triplets of residual related covariates, $X^g = (\mu^g, \sigma_g, R^g)$, using the following formula

$$\bar{R}^g_{\text{simulated}}(\Theta) = \bar{R}_{\text{theory}}(X^g, \Theta) + \epsilon^g\,, \qquad \epsilon^g \sim N(0, \sigma^2_{\bar{R}})\Big|_{\sigma_{\bar{R}}=1000} \tag{17}$$

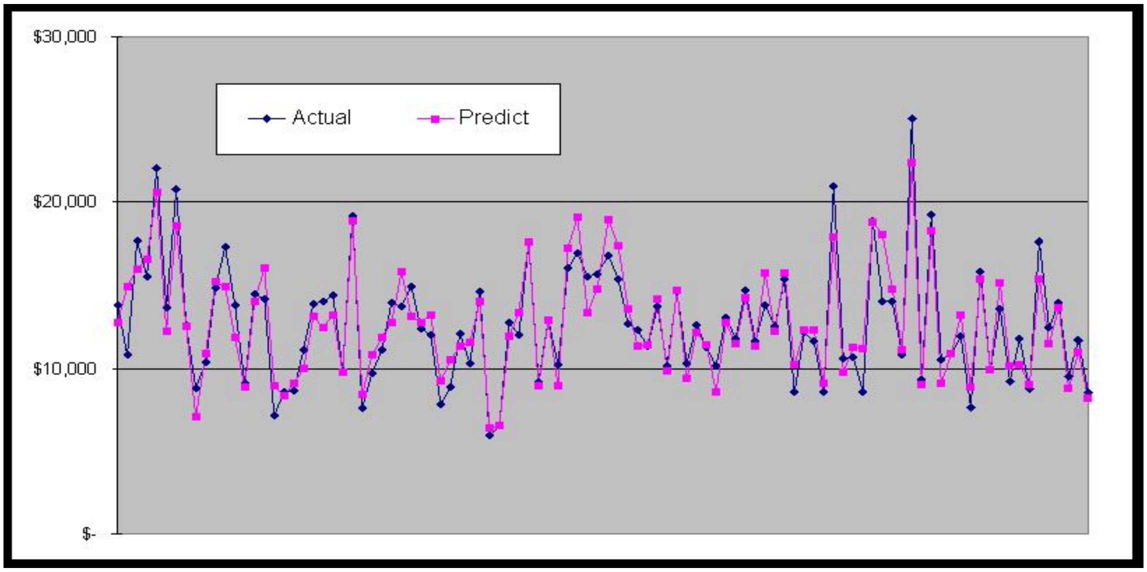

Figure 1: Nonlinear least square regression result using the simulated data

where the distributions of (μ^g, σ_g, R^g) are chosen to be close to the summary statistics of the real data at group level. The sample size is chosen to be sufficiently large (about three time as that of the real data) so as to avoid possible confounding due to small sample problem.

Table 1: *Summary statistics of the simulated data at the inter-group level*

	Mean	Volatility (%)
μ	0.02(/month)	13
σ^2	0.0001(/month)	—
R	16,301 ($)	13
$\bar{R}_{\text{simulated}}$	11,636 ($)	25

4.2 Nonlinear Least Square Estimation

As pointed out in the end of subsection 3.3 and shown in Table 2, there is severe co-linearity in the nonlinear least square estimation. Even with only three non-cost related parameters (α, ν and ρ) being estimated, the condition number for the Hessian matrix turns out to be rather high at 42.5. Fortunately, it is still possible to estimate these three parameters reasonably well, as can be seen from the table. The root mean square error and the values of the estimated parameters agree with the simulation input within errors. Figure 1 shows the overall regression quality, with the simulated data points in blue and fitted data points in magenta. The horizontal axis is in no particular order and the lines are simply to guide the eyes.

Table 2: *Regression results (with $\beta = 1000$ and $\kappa = 300$ fixed) using the simulated data*

RMSE	Condition Number	$\hat{\alpha}$	$\hat{\nu}$	$\hat{\rho}$
992 ($)	42.5	1770(75)	0.39(2)	0.014(4)

4.3 Sensitivity Analysis

Recall that the original estimation problem contains residual process related parameters (the total number is three times the number of groups) and decision-making related parameters (the total number is 6). These two sets of parameters can be treated separately by a two-step procedure. The major difficulty appears on estimating the 6 decision-making related parameters via only one nonlinear algebraic equation (10). Section 3.3 provides a foresight to an identification problem. However, we find that that ambiguity or more precisely model redundancy turns out insensitive to the quantities of interest. Namely the utility cash flow (value function) and the optimal stopping rule $\bar{R}$. We reveal these findings by sensitivity analysis.

If we were to include both cost parameters (β and κ) in the estimation, the co-linearity would completely ruin the regression. This is not to say that the exercise here is completely

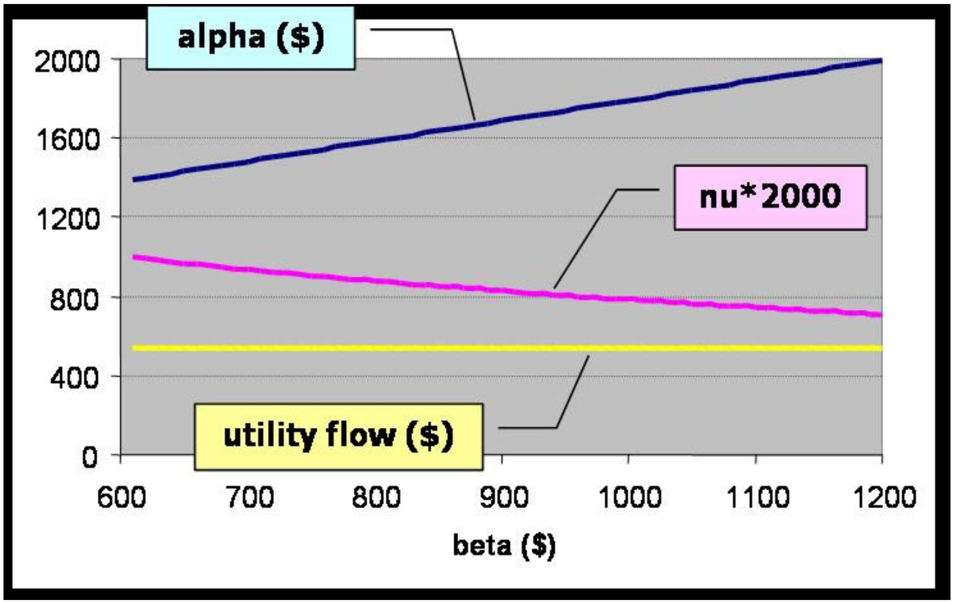

Figure 2: Regressed α (blue), ν (magenta), and the utility flow evaluated at the threshold (yellow) as functions of β. In order to put all three functions on the same graph, the value of ν is scaled by 2000.

useless. Rather, it speaks to the fact that there is redundancy in parameterizing the utility function for the rental company. The redundancy makes the optimal stopping problem insensitive to some of the parameters that we are trying to estimate. Additional information is necessary to resolve the redundancy.

In the absence of additional information, we can at least show that, while the estimated parameters depend on the assumed value of β and κ, the utility flow function and the option exercising threshold are essentially invariant in β and κ within reasonable ranges. In figure 4.3 the estimated $\hat{\alpha}$ (blue) and $\hat{\nu}$ (magenta) do change as β changes. However, the estimated utility flow function (yellow), as defined by $\hat{u}(r;\Theta) = \hat{\alpha}(r/R)^{\hat{\nu}} - \beta(r/R)^{-\hat{\nu}}$, is extremely flat in the entire range of $\beta \in (600, 1200)$. In figure 4.3 we depict the option exercising threshold $\bar{R}$ as a function of κ, while keeping everything else fixed. The figure indicates that the dependence is very weak. Even though κ changes by as much as a factor of 5, the corresponding change of $\bar{R}$ is less than 1%.

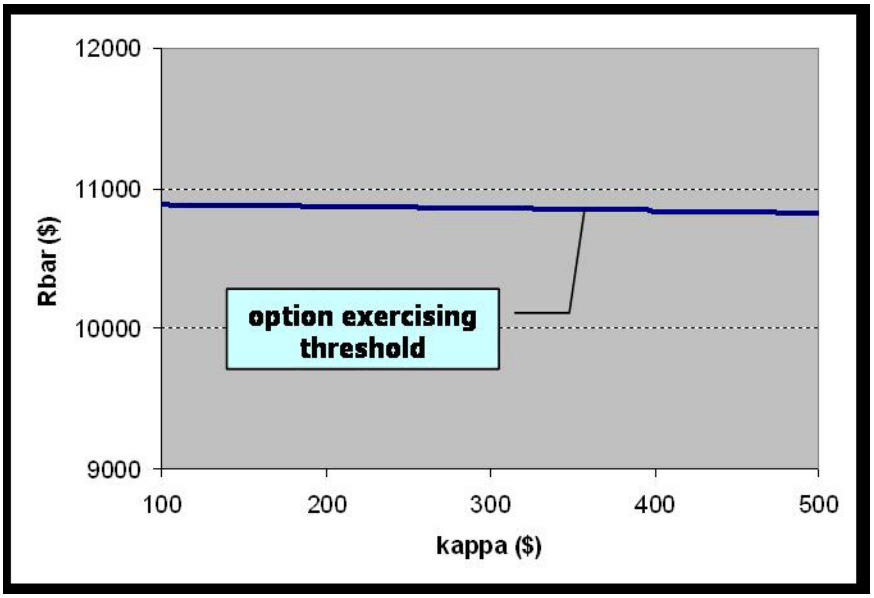

Figure 3: $\bar{R}$ as a function of κ

5 Conclusion

It is of great interest to understand the rental vehicle fleet planning for consumption capacity in tourism. Our study aims to retrieve a rental company's economic preference from its de-fleeting process.

We have developed a model based on the optimal stopping theory to describe the de-fleeting process for a vehicle rental company. Taking advantage of the pricing peculiarity in the industry, we also have introduced a specific econometric procedure to estimate the parameters in the model. We then used simulated data to illustrate how the estimation procedure could be applied to real data. Technical issues pertinent to the estimation procedure, such as the identifiability of model parameters and the redundancy in parameterizing the utility flow function, were elucidated.

While we have been concentrating on how to model the de-fleeting process for a rental company in this paper, the original motivation of the work is to study contract design between a vehicle OEM and a rental company. The utility flow function estimated using the methodology developed here serves as an important input to that study.

References

[1] A. K. Dixit and R. S. Pindyck, "Investment Under Uncertainty," Princeton University Press, 1993.

[2] W. H. Greene, "Econometric Analysis," Fifth Edition. Prentice Hall, 2003.

[3] J. E. Pachon, E. Iakovou, , C IP, and R. Aboudi. A synthesis of tactical fleet planning models for the car rental industry. IIE Transactions, 35(9), 907-916. 2003