Asia Pacific Journal of Information Systems
Vol. 29 No. 4 (December 2019), 673-689

ISSN 2288-5404 (Print) / ISSN 2288-6818 (Online)
https://doi.org/10.14329/apjis.2019.29.4.673

Smart Pricing in Action: The Case of Asset Pricing for a Rent-a-Car Company

Chang Hee Han[a], Seongmin Jeon[b,*], Sangchun Shim[c], Byungjoon Yoo[d]

[a] *Professor, College of Business and Economics, Hanyang University, Korea*
[b] *Associate Professor, College of Business, Gachon University, Korea*
[c] *Adjunct Professor, College of Business and Economics, Hanyang University, Korea*
[d] *Professor, College of Business Administration, Seoul National University, Korea*

ABSTRACT

The Internet enables businesses to acquire a great deal of information, including prices in the open markets. In this study, we investigate what the value of reference price information is to a company in the market and how the company can make use of such information. Using business analytics, we were able to estimate prices of used cars for a rent-a-car company. The results show that a smart pricing information system is useful for collecting online reference price information and for estimating future prices of used cars and rental prices.

Keywords: Smart Pricing System, Business Analytics, Rent-A-Car Industry, Electronic Markets, Mechanism Design

Ⅰ. Introduction

In traditional microeconomic theory, a state of perfect information is assumed in perfect competition. With the emergence of electronic markets on the Internet, almost perfect information is available to compare product offerings, including price. However, a considerable body of empirical literature on information systems demonstrates that price dispersion exists on the Internet. It proves that this information is not perfect, and that transactional friction still exists. In this paper, a multiple regression price estimation model for used cars is set up for a rent-a-car company, based on an analysis of a price data set available on the Internet. The rapid depreciation in used car prices or the "lemons problem" is well known (Akerlof, 1970). A reduced-form model, which is computationally feasible based on the reference prices, was delivered to the rent-a-car company to improve its current used car pricing systems.

Perfect competitive markets are those in which a large number of small buyers and sellers deal with

*Corresponding Author. E-mail: smjeon@gachon.ac.kr Tel: 82317505187

a homogeneous product, and a single small firm does not influence price allocation, instead it acts as a price taker (Mankiw and Taylor, 2006). In addition, in a perfectly competitive market, the mobility of the factors of production is perfect in the long run, and both the producers and consumers have perfect information regarding the product. Stigler suggests that, in an environment of price dispersion, information about market prices allows consumers to find lower prices for a given product or for horizontally differentiated substitutes (Stigler, 1961).

As a result of the proliferation of electronic markets via the Internet, there has been an expectation that frictionless commerce will emerge, wherein perfect information to compare product offerings will lead to higher competition and a subsequent erosion of profits (Brynjolfsson and Hitt, 2000). While lower friction exists in many dimensions of Internet competition, branding, awareness, and trust remain important sources of heterogeneity among Internet retailers. A number of studies have found analytical and empirical support for the presence of lower prices on the Internet relative to traditional channels (Brynjolfsson and Smith, 2000; Degeratu et al, 2000; Zettelmeyer et al., 2006).

Conversely, a group of researchers have found that higher prices (Bailey, 1998; Lal and Sarvary, 1999) and price dispersion (Chellappa et al., 2011) exist on the Internet. They argue that price dispersion exists because of the nature of e-retailers as well as the characteristics of the products they sell (Walter et al., 2006). Therefore, there is some evidence that, even in Internet-based markets, some of the frictions that mitigate head-on price-based competition will remain.

Predicting the price of used cars has recently become a popular topic of study as price information is readily available on the Internet. Listian (2009) points out that a regression model built using Support Vector Machines or SVM can predict the price of a car. Through multiple regression analysis, Richardson (2009) finds that hybrid cars retain their value for longer than traditional cars. He argues that car producers are producing cars that are comparatively more durable these days. Wu et al. (2009) conduct a car price prediction study by using a neuro-fuzzy knowledge-based system. They take into consideration the following attributes: brand, year of production, and type of engine. Their prediction model produced similar results as the simple regression model. Moreover, they created an expert system, named ODAV or Optimal Distribution of Auction Vehicles, to help meet the high demand by car dealers to sell cars at the end of the leasing year. A regression model based on a k-nearest neighbor machine learning algorithm is used to predict the price of a car. Gongqi et al. (2011) proposes a model built using ANN or Artificial Neural Networks to predict used car prices. His model reflects several attributes such as miles passed, estimated car life, and brand. The model differentiates itself from previous works as it can deal with nonlinear relations in data.

Furthermore, Pudaruth (2014) applies various machine learning algorithms, namely: k-nearest neighbors, multiple linear regression analysis, decision trees, and naïve Bayes. The data set used to create the prediction model is collected manually from local newspapers over a period of less than one month, as time can have a noticeable impact on the price of the car. His model considers such attributes as brand, model, cubic capacity, mileage in kilometers, production year, exterior color, transmission type, and price. However, the author finds that neither naïve Bayes nor decision tree is able to predict and classify numeric values. Additionally, because of the limited data set, they could not provide high classification performances. Noor and Jan (2017) build a model for car price prediction by using multiple linear regression. The data set is created during a two-month period and includes features like price, cubic capacity, exterior color, date when the ad was posted, number of ad views, power steering, mileage in kilometers, rim type, type of transmission, engine type, city, registered city, model, version, make, and model year. After applying the feature selection, the authors consider only engine type, price, model year, and model as input features. Even though many researchers have attempted to build a model to predict used car prices, we could not find many accurate or practical models with large volume data sets covering extended periods of time.

In this study, a multiple regression price estimation model for used cars is set up for a rent-a-car company, based on analysis of the price data set available on the Internet. The reduced-form model, which is computationally able to provide estimated price by making use of reference prices, is delivered to the rent-a-car company.

In this context, the two important research questions of the present study are as follows:

- What is the value of reference price information in the current open market to an enterprise that plans to estimate price levels in the future?
- What factors are influential in future price estimation?

To address these questions, we need to assess the impact of reference price information on pricing processes and the value of the information to the enterprise. However, as IT evolves from a mere productivity tool to a more pervasive and strategic business tool, the measurement of its value to an enterprise has become more challenging. A framework is needed to collect and make use of the reference price information.

We present a framework, based on an analysis of open-market data, for future price estimation by making use of reference prices available online. We apply our methodology to help a major rent-a-car company in Korea to systemize the estimation processes for used car pricing by using reference price information. Our results identify the factors that affect future price levels. We find that reference price information helps the company to estimate prices for used cars.

The remainder of this paper is organized as follows. Section 2 reviews related research. Section 3 introduces data and research contexts. Section 4 develops the mechanisms that form the basis of our methodology for reference price systems. Section 5 presents the results from empirical testing of a prototype that implements the methodology. Finally, section 6 discusses the implications of the results and concludes the paper.

Ⅱ. Literature Review

The literature reviewed for this study can be classified into categories of pricing in competitive and online markets, used car pricing, reference prices, and design mechanisms.

2.1. Pricing across Competitive and Online Markets

In the classic economic models, efficiency is maximized when all the welfare that enhances trades is executed. In retail markets, efficiency can be reached when prices are set equal to the marginal cost of the retailers. Marginal cost pricing can be an efficient outcome as pricing above the level of marginal cost excludes welfare enhancing trades from consumers. Economic theory predicts that high consumer search costs will lead to prices above marginal cost in equilibrium (Hotelling, 1929; Salop, 1979)

If electronic markets allow consumers to more easily determine retailer's prices and product offerings, these lower search costs will likely lead to lower prices for both homogeneous and differentiated goods (Bakos, 1997). More advantageous retailer cost structures may also contribute to lower price levels in electronic marketplaces.

Lee (1997) conducted one of the earliest studies involving pricing in electronic marketplaces. His study analyzes prices in both electronic and conventional auction markets for used cars. He finds that prices in electronic marketplaces are higher than in conventional markets and that this price gap between the two markets will increase over time. This finding seems opposed to the market efficiency hypothesis. However, the context of his study is an auction market where the good is sold to the bidder with the highest valuation. Higher prices in an auction market may be a signal of more efficient auction markets rather than a general retail market. Bailey (1997) provides another test of the efficiency hypothesis in electronic marketplaces. He examines Internet market efficiency by comparing the prices for books, CDs, and software sold on the Internet and through conventional channels. Bailey finds that electronic channels have higher prices than conventional channels. Bailey argues that the higher prices he observes could have been caused by market immaturity. Brynjolfsson and Smith (1999) analyze data on the prices for books and CDs sold through the Internet and conventional channels. They find that prices are lower on the Internet than in conventional outlets after considering costs from shipping and handling, delivery, and sales taxes.

One of the predictions related to the emergence of electronic markets is that price elasticity (i.e., the percent change in demand caused by a percent change in price) will be higher online than offline, because electronic markets enable consumers to search for information about competitive offerings at a lower cost (Alba et al., 1997; Smith et al., 2001). Price elasticity is useful for measuring how sensitive consumer demand is to price changes. Price elasticity could be an important signal of market efficiency. Consumers are more sensitive to small changes in prices in efficient markets. Higher price elasticity may result from the low search costs or low switching costs in electronic marketplaces. Goolsbee (1999) uses survey data to examine how sensitive customers are to sales tax rates. He finds that online consumers are very sensitive to tax policies: consumers who are subject to high sales taxes are more likely to purchase online. Evaluating products online could lead to missing information regarding the characteristics of the product (Degeratu et al., 1998; Lee, 2007), and consequently missing information could lead consumers to rely on other signals of quality, such as brand. Increased information about product characteristics and quality allows consumers to ascertain their valuation of a product with higher precision and thus to find a product that better fits their needs. All else being equal, product information is likely to make consumers less price sensitive, as it causes them to focus their search on product characteristics and quality rather than on price (Gupta et al., 2004). This assertion is founded on information integration theory (Anderson, 1968; Degeratu et al., 2000), which suggests that consumers assign importance weights and values to available search attributes and then tally them to make a purchase decision.

2.2. Context of Used Car Pricing

The well-known "lemons problem" refers to the rapid depreciation of used car values. Even though informational asymmetries may lead to market failure, pricing mechanisms exist in the secondary market for used cars. As a result of technological advances, modern vehicles are generally better maintained and easier to monitor than were vehicles at the time of Akerlof's classic study (Akerlof, 1970). In fact, empirical evidence on the severity of the lemons problem is quite mixed, and even studies that do find evidence supporting it find only weak effects (Emons et al., 2009; Genesove, 1993). More recent works by Engers et al. (2009) do not find any evidence supportive of a lemons problem. The presence of heterogeneous consumers results in strict gains from trade in the operation of a secondary market; however, as Rust (1987) shows, such a market is observationally equivalent to a homogeneous consumer market when an appropriate "representative consumer" is chosen. Konishi and Sandfort (2002) has extended this theory to account for transaction costs.

Researchers more often predict prices of products using previous data, so a multivariate regression model helps in classifying and predicting numeric values. Kuiper (2008) uses this model to predict the price of 2005 General Motor (GM) cars. Price prediction for cars does not require any special knowledge, so the data available online is sufficient to predict prices. Using this easily available data, he introduced variable selection techniques that helped in finding which variables are more relevant for inclusion in model. Similarly, Listiani (2009) uses SVM to predict the prices of leased cars. SVM handles high dimensional data well and avoids both under-fitting and over-fitting issues. However, the technique does not show why, in terms of variance and mean standard deviation, SVM is better than simple multiple regression.

Prieto et al. (2015) presents the results of expectation theory investigated in used car markets. A hedonic price model is developed to explain a price structure where consumers avoided risk when the used car's reliability is below the expected reference value and when it is above the expected reference value. The model shows how automobile quality affects residual values and how buyers evaluate used cars.

2.3. Customer Value Theory and Conceptual Framework

There are several different theories of reference pricing, which are usefully summarized and discussed by Briesch et al. (1997) and Kalyanaram and Winer (1995). Some of these theories have purely psychological bases and do not precisely specify an underlying process of reference price formation.

According to adaptation theory, individuals adopt a specific "reference point" and then take note of

<Table 1> The Definitions of Reference Price, Estimated Price, and Transaction Price

	Definition	Source
Reference Price	The price encountered previously or the price information available on open online marketplaces	Briesch et al., 1997; Kalyanaram and Winer, 1995
Estimated Price	The price estimated from the projection based on the reference price data set	Erdem et al., 2008; Hendel and Nevo, 2006; Rust, 1987; Sun, 2005
Transaction Price	The price at which a transaction occurs in the real world	Harless and Hoffer, 2002

changes from that point. Although this theory is plausible, as human behavior has been shown to exhibit this pattern in many contexts, it explains neither why consumers adopt a particular price level as a reference point nor how they update this price level once exposed to a change. The expectation process has been modeled structurally and shown to exist. If the true process is one of forward-looking price expectations (Erdem et al., 2008; Hendel and Nevo, 2006; Sun, 2005), then empirical estimation of a reduced form model, with a reference price variable based on past prices, will find that consumers care about reference prices because these prices serve as proxies for forward-looking price expectations, which are based on observations of past prices.

To clarify the concepts of reference price, estimated price, and transaction price, <Table 1> summarizes the definitions used in previous research and this study.

Ⅲ. Data and Research Context

Our data set is provided by one of the largest rent-a-car service providers in Korea. This organization pioneered IS-driven change initiatives in its industry and undertook a project to implement pricing systems based on business analytics. The company manages more than 60,000 cars and sells off used cars with proceeds amounting to nearly $20 million annually. The company's main revenue comes from long-term rent (2-3 years), which accounts for more than 85% of total rentals. Before beginning the aforementioned project, factors such as purchase price, maintenance cost, used car price, and insurance have become crucial not only to used car pricing but also to rental rates.

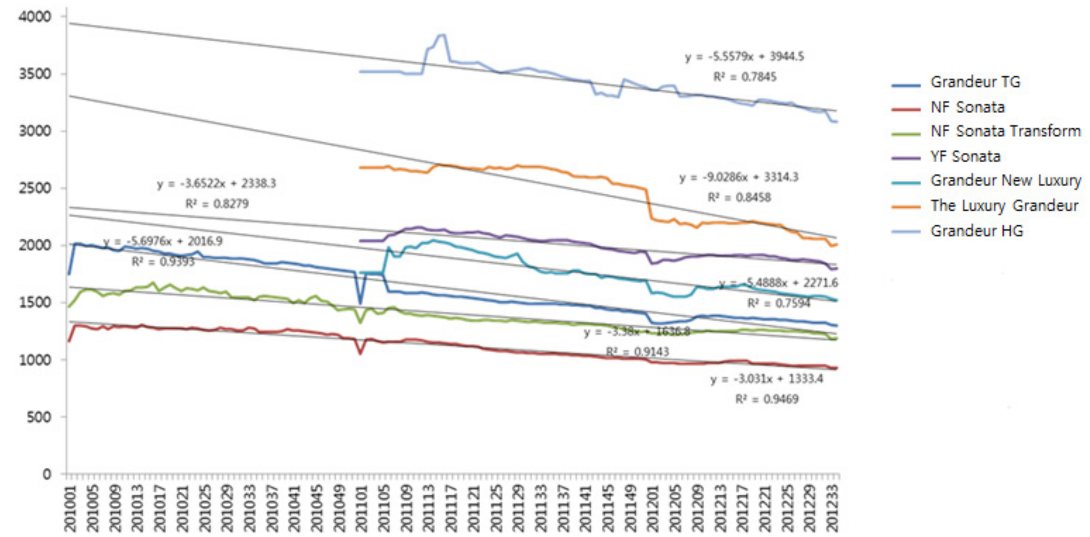
The research team used multiple data sources for this study. The first was a data set from one of the largest online open-market sites for used cars in Korea. The website provided the researchers with weekly transaction prices for used cars from January 2010 to August 2012. Unlike other websites for monthly used car prices, this site supplies weekly prices in a structured format. Thus, we were able to develop software programs to collect data from the website. Some of the data collected in this manner required filtering as some of the fields were inconsistent.

3.1. Data Description

Pricing projections were made for 15 out of 43 types of used cars by using multiple regression models modified for each type. For this purpose, we used the weekly transaction data collected from the open market site for used cars. The website, which was established in 2001, receives more than 100,000 visitors and information on more than 3,000 car registrations per day. It provides information for each used car (including passenger cars, buses, and trucks) with details including car descriptions, pictures, inspection results, and prices on the site. The website is updated with input from the staff in the nationwide branch offices. The search functions provided by the website were useful for collecting results regarding the current prices for used cars.

3.2. Depreciation over Time

It is obvious that the values of used cars will decrease with time. One approach to examining the decay of assets is to study the cost of using machines as they age. That is, as depreciation values can be estimated from changes in rental prices, the rent-a-car company will identify the pricing information over time from a publicly available source, which is very



<Figure 1> Price Changes over Time (Selective Car Types)

useful both in estimating the remaining value of the cars and in making decisions regarding their rental prices. One of the most crucial factors for determining the current values of used cars, is the time elapsed from the time of purchase. <Figure 1> shows the price changes for various car types over time. The single variable of elapsed time explains a considerable portion of the price changes in the majority of cases.

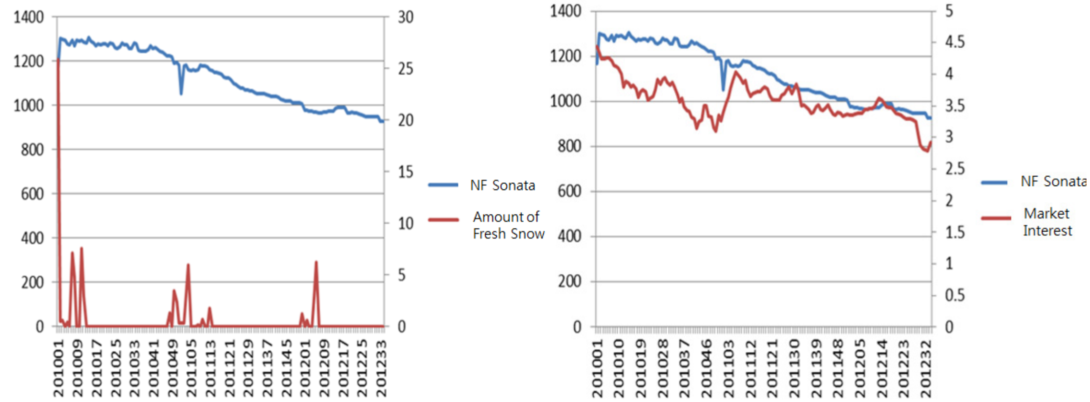
The price-change plots overlapped with the OLS regression line over time, and we find that all the slopes are obviously negative. The single variable used here has very high explanatory power. When we look at the R-square values, which vary from 0.7594 to 0.9469, we can see that the percentage of variation in the response variable that is explained by the model is higher than 80% in most cases.

The second set of data we utilize is the control data set collected from the Internet, including variables such as macroeconomic and weather conditions. Seasonal changes and external shocks will be controlled with these variables. The bank of Korea and Korea Meteorological Services provide a number of indexes for the period under investigation.

The weekly changes for macroeconomic variables are obtained from the Bank of Korea's public website (http://ecos.bok.or.kr). Since weather is one of the distinctive variables that influence used car transactions[1], additional variables are gathered from the open website of the Korean Meteorological Services (http://www.kma.go.kr/). After investigating the correlations among the variables in these data sets, we find that interest rates, exchanges rates, and consumer price indexes are important. Furthermore, weekly and monthly dummy variables are suitable to control weather variances in the interest of parsimony. <Figure 2> shows the amount of fresh snow and market share over time overlapped with a car price.

The third data set was collected from the company's 15,000 real transactions over the course of a year. The SK encar (http://www.encar.com) offers on-line

1) 10 Best Times to Buy a Used Car https://cars.usnews.com/cars-trucks/best-times-to-buy-a-used-car



<Figure 2> Car Price Overlapped with Amount of Fresh Snow and Market Interest over Time

<Figure 3> Scatter Diagrams by Prices and Time periods Across Car Types

used car advertising, vehicle history and pricing information services. The company was founded in 2014 and is based in South Korea. The record shows that the company primarily sold off several specific product groups, types of cars, due to the characteristics of the rent-a-car business. More than 50% of the total number of cars sold belong to the top five product groups. Specifically, some of compact car groups have less than 300 samples, with many deviations based on models and options. Due to the limited data availability, we selected a sample of car models with a statistically significant size to verify the estimation model. However, even these limited data provided extremely interesting insights into the need to consider reference prices in used car pricing. <Figure 3> illustrates the transaction price changes of major car models over time. We find that Hyundai Sonata and Grandeur, which form the main models, provide more samples than Kia Morning and Pride do.

A typical set of data for a car type is described below in <Table 2>. *Price (10,000 KRW)* represents a car price when the car is sold at the unit of 10,000 Korean Won. The variable of *idx_year_week* controls for weekly market changes. *y2005, y2006, and y2007* are the dummy variables equal to 1 if the car is manufactured in the year of 2005, 2006, and 2007, 0 otherwise. The reason we have only three year dummy variables is that rent-a-car company sells out all the three or older cars. *interest_rate, e_dollar, e_yen, e_euro,* and *CPI* are the macroeconomic variables representing interest rate, Won/US Dollar exchange rate, Won/Japanese Yen exchange rate, Won/Euro exchange rate, and Consumer Price Index. *m1* to *m12* are the control variables that capture

<Table 2> Descriptive Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Price (10,000 KRW)	551	1109.92	146.34	848.57	1420
idx_year_week	551	70.23	40.24	1	139
y2005	551	0.25	0.43	0	1
y2006	551	0.25	0.43	0	1
y2007	551	0.25	0.43	0	1
interest_rate	547	3.59	0.31	2.79	4.44
e_dollar	547	1134.43	37.29	1052.66	1242.58
e_yen	547	1374.21	77.76	1183.81	1547.43
e_euro	547	1520.22	50.92	1389.92	1673.46
CPI	547	102.98	2.56	98.80	106.20
m1	551	0.10	0.30	0	1
m2	551	0.09	0.28	0	1
m3	551	0.09	0.28	0	1
m4	551	0.09	0.28	0	1
m5	551	0.09	0.28	0	1
m6	551	0.09	0.28	0	1
m7	551	0.11	0.31	0	1
m8	551	0.11	0.31	0	1
m9	551	0.06	0.23	0	1
m10	551	0.06	0.23	0	1
m11	551	0.06	0.23	0	1

the effects of monthly changes within a year.

The price for this type of car varies from 8,480,000 KRW or Korean Won (about USD 8,000) to 14,200,000 KRW (about USD 13,000). This type of car was manufactured from 2005 to 2007; dummy variables are included for the years of manufacture. The macroeconomic variables of interest rate, exchange rate, and consumer price index are presented. The variables e_dollar, e_yen, and e_euro represent the exchange rates between the Korean Won and, respectively, the US Dollar, Japanese Yen, and EU Euro. Finally, monthly control variables are added as dummy variables to control monthly or seasonal changes.

Ⅳ. Model

In order to set up the estimation model, we consider the explanatory power achieved by the variables from the data sets. As shown in our previous study, the multiple regression methodology is utilized to identify the appropriate price level for cars. We build a linear combination of explanatory variables that explains the response variable of reference price. As a single "best" model may not exist for these data, we set up basic models and then modify them to increase the explanatory power of the multiple regression models. Often, when the goal of developing a multiple regression model is description or prediction, the primary issue that arises is determining which variables to include or exclude from the model. It is possible to include all potential explanatory variables in a regression model, but doing so often results in a cumbersome model that is difficult to understand. On the other hand, a model that includes only one or two explanatory variables may be much less accurate than a more complex model. Including redundant or unnecessary variables, not only creates an unwieldy model, but can also lead to less reliable test statistics and conclusions from corresponding hypothesis tests. If explanatory variables are highly correlated, then their effects in the model will be estimated with less precision, leading to larger standard errors and potentially to insignificant test results for individual variables that may be important in the model.

Thus, in this study, we again use the multiple regression models for 43 types of cars in the data set. In our empirical models, the dependent variable is the price in year t, and the independent variables are the cumulative distance travelled for the same period of time t.

Our multiple regression equation for used car pricing is

$$Y_i = \beta_0 + \beta_1(Car\ Info._i) + \beta_2(Car\ Type_i) + \beta_3(Time\ Control_i) + \beta_4(MacroEconomy_i) + \varepsilon_i \quad (1)$$

where $Car\ Info._i$ contains variables representing information about car *i*, such as mileage, manufactured year, new car price, accident history, options, and fuel type; $Car\ Type_i$ denotes the car type to which car *i* belongs; $Time\ Control_i$ represents the seasonal variables regarding the time at which car *i* is sold; and $MacroEconomy_i$ contains macroeconomic variables like exchange rate, income, interest rate, and CPI.

Next, we test the equation with the variables specified above to determine whether we have all of the necessary variables for the estimation. We eliminate variables with multi-collinearity issues resulting in high correlations among the variables, because we do not wish to take the "kitchen sink" approach of using many different options or methods to solve a problem or achieve a goal.

Ⅴ. Results

We identify the variables that are universally, statistically significant across all car types. The variables of week dummy, year dummy, interest rate, Korean Won/Dollar/Yen/Euro exchange rates, and monthly dummy are extracted from the whole set of variables. The multiple regression price estimation shows that the manufactured year, interest rate, exchange rate, consumer price index, and the month in which the car was sold are statistically significant in influencing the price.

One of the results for a car type is presented in <Table 3>. We first include all the identified variables in the regression model and then exclude the variables that are not statistically significant.

Though our model may have endogeneity and identification issues because the explanatory variables are highly correlated, we have nonetheless built a model with strong explanatory power, with R-squared values higher than 0.96. This may be interpreted as the model has overfitting issues. However, our concern is more about causal identification rather than practical estimation of used car prices. As we focus on implanting the price prediction, we have included every significant variable in our model (Verboven, 2002; Verboven and Brenkers, 2006).

According to our OLS regressions, week dummy, manufactured year, Dollar and Yen exchange rates, and CPI are estimated to be statistically significant in used car pricing equations, whereas interest rate and Euro exchange rate are statistically insignificant. Some of the monthly dummy variables are statistically significant, while others are not. Although we should

<Table 3> Regression Results

	(1)	(2)	(3)
idx_year_week	-1.388*** (0.273)	-1.104*** (0.196)	-1.209*** (0.178)
y2005	-3.242 (3.293)		
y2006	72.40*** (3.293)	74.02*** (2.854)	74.03*** (2.855)
y2007	179.3*** (3.293)	181.0*** (2.854)	181.0*** (2.855)
interest_rate	-11.25 (11.270)		
e_dollar	-0.525*** (0.079)	-0.521*** (0.069)	-0.479*** (0.061)
e_yen	0.428*** (0.061)	0.419*** (0.055)	0.387*** (0.049)
e_euro	-0.0207 (0.049)		
CPI	-39.07*** (4.516)	-42.21*** (3.700)	-39.76*** (3.156)

<Table 3> Regression Results(Cont.)

	(1)	(2)	(3)
m1	6.247 (6.920)	7.463 (5.889)	
m2	29.13*** (7.660)	29.78*** (6.403)	24.78*** (5.044)
m3	47.90*** (7.567)	48.95*** (7.053)	43.38*** (5.522)
m4	64.60*** (8.306)	65.57*** (7.969)	59.42*** (6.324)
m5	47.69*** (7.676)	49.62*** (7.069)	44.27*** (5.673)
m6	34.52*** (6.834)	36.56*** (6.393)	31.80*** (5.173)
m7	29.15*** (7.001)	29.86*** (5.592)	25.90*** (4.638)
m8	22.23*** (6.245)	25.25*** (5.155)	21.94*** (4.449)
m9	27.14*** (6.989)	32.70*** (6.331)	28.86*** (5.560)
m10	18.62*** (6.889)	23.14*** (6.045)	20.53*** (5.688)
m11	-7.574 (6.744)		
Constant	5,220*** (454.200)	5,456*** (384.200)	5,212*** (332.700)
Observations	547	547	547
R-squared	0.966	0.966	0.966

Note: The dependent variable represents the transaction prices of used cars.
The standard errors are reported in parentheses.
* Significant at the 10% level. ** Significant at the 5% level. *** Significant at the 1% level.

consider other variables such as major redesigns of cars, the weekly dummy is statistically significant, demonstrating that elapsed time is one of the most important factors in determining the prices of used cars. After obtaining the results from the model, we compare the estimated prices with the actual transaction data from the rent-a-car company. Most of the projected prices fall within the boundary of 5% gaps, which the industry team finds very useful.

The results reveal that, among the car-related information, manufactured year has a strong effect on prices. At the same time, fuel type is not correlated with used car prices. Further, the results illustrate that seasonal variables during the year have a signifi cant impact. Moreover, the model illustrates that macroeconomic variables, such as exchange rate, and CPI have a positive and significant effect on price, while interest rate does not.

Ⅵ. Conclusions and Discussions

In this paper we have presented a prediction model to valuate used cars and identified the key factors in the determination of prices in the used car market. We have used a multiple regression model utilizing the massive data available on the Internet. The most important finding of this study is that the reference price information available in the online open market is useful for estimating the prices of used cars. Although multiple regression equations are derived empirically, there is little doubt that they have generated significantly more accurate forecasts than those determined on the basis of historic transactions.

The estimation model of the multiple regressions confirms that the time variable, manufactured year, some of the exchange rates, and CPI are statistically significant in used car pricing equations. We argue that the interest rate and monthly dummy variables control for monthly or seasonal changes. This complements previous findings regarding household intentions to replace old cars (Marell et al., 2004) and hedonic price analysis (Prieto et al., 2015). In particular, our findings confirm that manufactured year has a strong significant impact on price. Unexpectedly, accident history has no statistically significant effects in our study. This may be due to variances in accident damage as some of the cars will be fine, with only minor damage, while others will lose much value due to major damage. The record of an accident saved in the vehicle history report may not provide quantitative information that is meaningful to buyers.

Moreover, it has been shown that exchange rates influence prices of used cars. This is because exports of used cars are growing in tandem with the rising popularity of and confidence in Korean vehicles overseas. According to a daily newspaper article[2], Russia and Jordan are two of the biggest importers.

The results of our study lead to several strategic marketing recommendations for used car retailers and augment previous literature (Parment, 2008). We can observe that elapsed time has a significant impact on prices. However, fuel type (gasoline, diesel, or LPG) has no significance, even though this fuel type likely moderates the high mileage of cars. As we have growing concerns about air pollution and environmental issues, the variable of fuel type should be looked over in the long run as the government has strengthened regulations related to CO_2 emissions tax (Carling et al., 2013). Finally, professional sellers should communicate more clearly online as all transaction evidence will be stored in the server.

Ⅶ. Limitations and Future Studies

Important issues await further investigation. Some estimation issues remain that were not considered in this study. It would be interesting to complement our study in at least two ways. First, by performing a larger study including more car models and wider geography than the ones examined here. Furthermore, if additional information had been provided about the omitted variables, such as major model changes or the usage behavior patterns of previous owners, a more fruitful discussion may have resulted. However, it has been demonstrated that the multiple

2) Used Korean cars hot property overseas http://koreajoongangdaily.joins.com/news/article/article.aspx?aid=2951564

regression equations appear to generate highly accurate forecasts of used car prices for the 46 car types included in the study. Thus, we expect the industry to continue to make use of the multiple regression equations for used car pricing in the future.

Second, the influence of used car transactions on the new car market could be investigated, especially in the light of new technology. In particular, electric cars are becoming increasingly popular in many countries. In the future, the diffusion of electric vehicles may change the price patterns of used cars.

This study does not consider the demand and supply in the entire used car market, but rather it focuses on the effects of the reference price information that is published in the market on a single company. The question of whether online price reference systems are useful in a company's price estimation is raised in this study. The rent-a-car company is able to collect the data set from the largest B2C website for used cars and use it as a reference for their used car pricing. A drawback associated with this approach may be that some car types have an insufficient number of samples to produce statistically significant results. Furthermore, outstanding outliers exist; thus, the operator of the estimation systems should manually verify the quality of the data set in the systems.

It is intuitively obvious that the proposed estimation systems are effective and are subject to empirical evaluation in the future. Additional monitoring is required with regard to the quality of the estimation systems. At present, the rent-a-car company finds the estimation systems to be very useful smart services and applies the estimates generated by the systems in their pricing decisions.

Acknowledgement

The authors are thankful for the support by the Institute of Management Research at Seoul National University.

<References>

[1] Akerlof, G. (1970). The market for lemons: Qualitative uncertainty and the market mechanism. *Quarterly Journal of Economics, 84*(3), pp. 488-500.

[2] Anderson, N. H. (1968). A simple model for information integration. R. P. Abelson, E. Aronson, W. J. McGuire, T. M. Newcomb, M. J. Rosenberg, P. H. Tannenbaum, eds. *Theories of Cognitive Consistency: A Sourcebook*. Rand McNally, Chicago, 731-758.

[3] Alba, J. W., Lynch, J., Weitz, B., Janiszewski, C., Lutz, R., Sawyer, A., and Wood, S. (1997). Interactive home shopping: Consumer, retailer, and manufacturer incentives to participate in electronic marketplaces. *Journal of Marketing, 61*(3) 38-53.

[4] Bakos, J. Y. (1997). Reducing buyer search costs: implications for electronic marketplaces. *Management Science, 43*(12), 1609-175.

[5] Bailey, J. P. (1998). *Intermediation and electronic markets: Aggregation and pricing in Internet commerce.* Ph.D. Thesis, Technology, Management and Policy, Massachusetts Institute of Technology, Cambridge, MA.

[6] Bailey, J. P. (1998). *Electronic commerce: prices and consumer issues for three products: Books, compact discs, and software.* Organisation for Economic Co-Operation and Development, OCDE/GD (98), 4.

[7] Briesch, R. A., Krishnamurthi, L., Mazumdar, T., and Raj, S. P. (1997). A comparative analysis of reference price models. *Journal of Consumer Research, 24*(2), 202-214.

[8] Brynjolfsson, E., and Hitt, L. M. (2000). Beyond computation: Information technology, organizational

transformation and business performance. *Journal of Economic Perspectives, 14*(4), 23-48.
[9] Brynjolfsson, E., and Smith, M. D. (1999). Frictionless commerce? A comparison of internet and conventional retailers. *Working Paper*.
[10] Brynjolfsson, E., and Smith, M. D. (2000). Frictionless commerce? A comparison of internet and conventional retailers. *Management Science, 46*(4), 563-585.
[11] Carling, K., Hakansson, J., and Jia, T. (2013). Out-of-town shopping and its induced CO2-emissions. *Journal of Retailing and Consumer Services, 20*, 382-388.
[12] Chellappa, R. K., Sin, R. G., and Siddarth, S. (2011). Price formats as a source of price dispersion: A study of online and offline prices in the domestic US airline markets. *Information Systems Research, 22*(1), 83-98.
[13] Degeratu, A., Rangaswamy, A., and Wu, J. (1998). Consumer choice behavior in online and regular stores: The effects of brand name, price, and other search attributes. Presented at *Marketing Science and the Internet, INFORM College on Marketing MiniConference*. Cambridge, MA. 6-8 March.
[14] Degeratu, A., Rangaswamy, A., and Wu, J. (2000). Consumer choice behavior in online and traditional supermarkets: The effects of brand name, price, and other search attributes. *Internat. J. Res. Marketing, 17*(1), 55-78.
[15] Emons, W., and Sheldon, G. (2009). The market for used cars: new evidence of the lemons phenomenon. *Applied Economics, 41*(22), 2867-2885.
[16] Engers, M., Hartmann, M., and Stern, S. (2009). Are lemons really hot potatoes? *International Journal of Industrial Organization, 27*(2), 250-263.
[17] Erdem, T., Keane, M., and Sun, B. (2008). A dynamic model of brand choice when price and advertising signal product quality. *Marketing Science, 27*(6), 1111-1125.
[18] Genesove, D. (1993). Adverse selection in the wholesale used car market. *Journal of Political Economy*, 101(4), 644-665.
[19] Gongqi, S., Yansong, W., and Qiang, Z. (2011). New model for residual value prediction of the used car based on BP neural network and nonlinear curve fit. In *Measuring Technology and Mechatronics Automation (ICMTMA), 2011 Third International Conference on* (Vol. 2, pp. 682-685). IEEE.
[20] Goolsbee, A. (1999). In a world without borders: the impact of taxes on internet commerce. *Working Paper*, University of Chicago.
[21] Gupta, A., Su, B., and Walter, Z. (2004). An empirical study of consumer switching from traditional to electronic channel: A purchase decision process perspective. *International Journal of Electronic Commerce, 8*(3) 131-161.
[22] Harless, D. W., and Hoffer, G. E. (2002). Do women pay more for new vehicles? Evidence from transaction price data. *The American Economic Review, 92*(1), 270-279.
[23] Hendel, I., and Nevo, A. (2006). Measuring the implications of sales and consumer inventory behavior. *Econometrica, 74*(6), 1637-1673.
[24] Hotelling, H. (1929). Stability in Competition. *The Economic Journal*, 41-57.
[25] Kalyanaram, G., and Winer, R. S. (1995). Empirical generalizations from reference price research. *Marketing Science, 14*(3, Part 2), G161-G169.
[26] Konishi, H., and Sandfort, M. T. (2002). Existence of stationary equilibrium in the markets for new and used durable goods. *Journal of Economic Dynamics and Control, 26*(6), 1029-1052.
[27] Kuiper, S. (2008). Introduction to multiple regression: How much is your car worth? *Journal of Statistics Education, 16*(3).
[28] Lal, R., and Sarvary, M. (1999). When and how is the internet likely to decrease price competition? *Marketing Science, 18*(4), 485-503.
[29] Lee, D.W. (2007). Understanding price adjustments in E-commerce. *Asia Pacific Journal of Information Systems, 17*(4), 113-132.
[30] Lee, H. (1997). Do electronic marketplaces lower the price of goods. Communications of the ACM, 41(12).

[31] Listiani M. (2009). *Support vector regression analysis for price prediction in a car leasing application*. Master Thesis. Hamburg University of Technology

[32] Mankiw, M. G., and Taylor, M. P. (2006). *Microeconomics: Thomson Learning*.

[33] Marell, A., Davidsson, P., Garling, T., and Laitila, T. (2004). Direct and indirect effects on households' intentions to replace the old car. *Journal of Retailing and Consumer Services, 11*, 1-8.

[34] Noor, K., and Jan, S. (2017). Vehicle price prediction system using machine learning techniques. *International Journal of Computer Applications, 167*(9), 27-31.

[35] Parment, A., (2008). Distribution strategies for volume and premium brands in highly competitive consumer markets. *Journal of Retailing and Consumer Services, 15*, 250-265.

[36] Prieto, M., Barbara, C., and George, B. (2015). Using a hedonic price model to test prospect theory assertions: The asymmetrical and nonlinear effect of reliability on used car prices. *Journal of Retailing and Consumer Services, 22*, 206-212.

[37] Pudaruth, S. (2014). Predicting the price of used cars using machine learning techniques. *International Journal of Computer and Information Technology, 4*(7), 753-764.

[38] Richardson, M. S. (2009). *Determinants of used car resale value*. Retrieved from: https://digitalcc.coloradocollege.edu/islandora/object/coccc%3A1346

[39] Rust, J. (1987). Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher. *Econometrica: Journal of the Econometric Society, 55*(5), 999-1033.

[40] Salop, S. (1979). Monopolistic competition with outside goods. *Bell Journal of Economics, 10*, 141-156.

[41] Smith, M. D., Bailey, J., and Brynjolfsson, E. (2001). Understanding digital markets: Review and assessment. *Working Paper* 4211-01, Sloan School of Management, MIT, Cambridge, MA.

[42] Stigler, G. J. (1961). The economics of information. *The Journal of Political Economy, 69*(3), 213-225.

[43] Sun, B. (2005). Promotion effect on endogenous consumption. *Marketing Science, 24*(3), 430-443.

[44] Verboven, F. (2002). *Quantitative study to define the relevant market in the passenger car sector* (Report for the European Commission for competition).

[45] Verboven, F., and Brenkers, R. (2006). Market definition with differentiated products: Lessons from the car market. In: Choi, J.P. (Ed.), *Recent Developments in Antitrust: Theory and Evidence*. MIT Press, Cambridge.

[46] Walter, Z., Gupta, A., and Su, B. C. (2006). The sources of on-line price dispersion across product types: An integrative view of on-line search costs and price premiums. *International Journal of Electronic Commerce, 11*(1), 37-62.

[47] Wu, J. D., Hsu, C. C., and Chen, H. C. (2009). An expert system of price forecasting for used cars using adaptive neuro-fuzzy inference. *Expert Systems with Applications, 36*(4), 7809-7817.

[48] Zettelmeyer, F., Morton, F. S., and Silva-Risso, J. (2006). *Scarcity rents in car retailing: Evidence from inventory fluctuations at dealerships* (No. w12177). National Bureau of Economic Research.

◆ About the Authors ◆

Changhee Han

Changhee Han is a professor at School of Business Administration in Hanyang University. His research interests are in analysis of ICT business models, development of management consulting methodology, and economic and policy issues of electronic commerce, etc. He has published on these topics in journals such as *Information Systems and e-Business Management, International Journal of Production Economics, European Journal of Operational Research, and Journal of Operations Research Society.*

Seongmin Jeon

Seongmin Jeon is an associate professor of Information Systems at College of Business, Gachon University. He has major areas of expertise for the use of data science methods to extract, process, and analyze large volume data. His works have been published in peer-reviewed international journals including *Journal of Interactive Marketing, Information & Management,* and *Electronic Commerce Research and Applications.*

Sangchun Shim

Sangchun Shim received Ph.D. in Business Administration at College of Business and Economics, Hanyang University, and MBA at George Washington University in USA.
He used to work for KT (Korea Telecom) and KT Rental (Present, Lotte Rental), and to lead a number of management proejcts using business analytics.

Byungjoon Yoo

Byungjoon Yoo is Professor at the College of Business Administration at Seoul National University. His research interests are in online marketplaces and management strategies of digital goods such as software products, online music songs and online games. He has published on these topics in journals such as *Management Science, the Journal of Management Information Systems, Decision Support Systems* and *the International Journal of Electronic Commerce.*

Submitted: March 19, 2019; 1st Revision: June 17, 2019; 2st Revision: August 25, 2019; Accepted: August 29, 2019