공학석사학위논문

Personalized pricing strategy under inaccurate prediction performance of reservation price

2020 년 8 월

서울대학교 대학원

산업공학과

강 영 은

Personalized pricing strategy under inaccurate prediction performance of reservation price

지도교수 박 용 태

이 논문을 공학석사 학위논문으로 제출함

2020 년 6 월

서울대학교 대학원

산업공학과

강 영 은

강영은의 공학석사 학위논문을 인준함

2020 년 6 월

위 원 장	이 경 식	(인)
부위원장	박 용 태	(인)
위 원	이 덕 주	(인)

Abstract

Personalized pricing strategy under inaccurate prediction performance of reservation price

Youngeun Kang
Department of Industrial Engineering
The Graduate School
Seoul National University

Due to the development of the digital economy, personalized pricing is getting a lot of attention from both companies and academia. Unlike the past, development of data collection environment, advanced data analysis algorithms, and hardware have made it easier to comprehend the preference of individual consumers, giving the opportunity to companies to predict reservation price of customers and set personalized price to them. Companies are interested in personalized pricing to expand their markets and gain more profit. However, it is impossible to accurately predict the reservation price, and it could cause adverse effects if incorrect prices are presented to consumers due to inaccurate reservation price prediction.

Within this context, this study establishes an agent-based model to examine the effect of the personalized pricing strategy according to the prediction accuracy of the reservation price. First of all, the utility function of the consumer including the price is constructed by collecting web data including price, consumer purchase decision, and behavior data. Next, the reservation price of the consumer is derived using the constructed utility function. Finally, an agent-based model considering prediction performance under a personalized pricing strategy is constructed and simulated. The agent-based model is adopted since it has the ability to investigate the effects of prediction performance, reflect customer heterogeneity and interactions under personalized pricing situation.

This study offers some important insights into product diffusion under the personalized pricing strategy. First, it provides insights on companies to maximize profits by analyzing the effects of inaccurate prediction situation. Second, it contributes to the existing diffusion research by constructing an agent-based model that includes social influence about personalized prices that were not considered under the existing uniform pricing strategy but should be considered under the personalized pricing strategy. Third, it covers the entire process including data preprocessing, utility function derivation, model establishment, and strategic analysis using actual web data.

Keywords: Personalized pricing, Reservation price, Prediction performance, Diffusion, Agent-based model

Student Number: 2017-28225

Contents

List of Tables

List of Figures

Chapter 1

Introduction

With the advent of the digital age, new business opportunities are emerging. In addition to the information provided by consumers, companies are able to collect more data about consumers by tracing their digital footprint on the Internet (Koen, 2014). With these changes, advanced algorithms and hardware were developed to analyze the collected data. In this context, companies are trying to gain insights about consumers by analyzing the data and link it to their business. For example, web analytics tools such as Google Analytics help firms categorize their customers and increase marketing effectiveness through automatized reports about characteristics of the users who visit their website and observed behavior within the site.

Given the advancement in customer data analysis, personalized pricing has received great attention in both academia and practice. Personalized pricing is a pricing strategy that charges different prices for the same product/service according to the preferences of individual consumers. In the past, it was difficult to grasp the preference of individual customers because of the difficulty in obtaining and analyzing data about consumers. Nowadays, with the development of data collection and analysis, it becomes easy to draw the preference of individual customers which makes personalized pricing possible. As an example, Uber charges their consumers by introducing a machine learning model that takes into account a number of variables such as the consumer's boarding history, weather, holidays, time zone, and traffic (Martin, 2019). Personalized pricing is an important topic that attracts attention from companies, consumers, and policy researchers because it has multiple effects like the potential to profit maximization, market size expansion, consumer trust decline, and privacy infringement.

The key to this personalized pricing strategy is utilizing data to identify consumers' heterogeneous preferences for goods and services, enabling firms to identify the highest price that an individual consumer would like to pay for the product and present the price to the consumer. However, the problem is that it is impossible to accurately predict the reservation price because of the limitations to the customer data accessibility, analysis technique, and consumer privacy protection behavior.

Conventional approaches that deal with the effects of personalized pricing strategy have been subject to mathematical models, not considering the interaction of heterogeneous consumers under the personalized pricing strategy and assuming reservation price is able to be identified under the availability of consumer data (Chen, Choe, & Matsushima, 2020; Montes, Sand-Zantman, & Valletti, 2019). Consumer interaction is one of the main factors that forms consumers' perception which is the biggest obstacle to the introduction of personalized pricing strategies. Therefore, including the consumers' interaction process is crucial to clearly understand the ripple effects of the strategy. Moreover, it is important to consider the prediction performance of reservation prices because companies have to consider the impact of inaccurately determined reservation prices and take different actions based on it. For example, in the case of inaccurate reservation price prediction, one of the possible strategies is pricing lower than the expected reservation price to prevent consumer churn.

In this context, this study aims to analyze the adoption behavior of consumers by creating an agent-based model for personalized pricing considering the inaccurate reservation price prediction situation. Agent-based modeling is a methodology that is widely used in the field of technology diffusion, which has the advantage of considering the heterogeneity and interaction of consumers. In particular, considering the fact that personalized prices are formed from individual preferences, this agent-based modeling method is suitable because it is able to model heterogeneous customer characteristics and their interaction. The procedure for creating the agent-based model of personalized pricing is as follows. First, the utility function of the consumer which shows the relationship between price, customer behavior, and purchase action is constructed by utilizing web data. Next, the reservation price of the consumer is derived using the established utility function. At last, we create and simulate an agent-based model that includes a decision-making process for individual consumers.

The contribution of this study is as follows. First, the utility function of the consumer is constructed and the reservation price for the product is derived by using real data generated on the web. Second, an agent-based model is constructed that includes social influence about individual consumer prices that were not considered in the existing uniform pricing strategy but should be considered in the personalized pricing situation. Third, the concept of prediction performance is subdivided into accuracy and precision to examine the impact and provided insight into the behavior of companies to maximize profits.

The remainder of this paper is organized as below. First, the concepts and methodologies from the proposed approach are introduced in Chapter 2. Next, Chapter 3 describes the proposed model to consider prediction performance in the personalized pricing strategy. Then, Chapter 4 explains a case study that utilizes real e-commerce data. Chapter 5 provides the experiment design and simulation results. Finally, in Chapter 6, the paper ends with discussions and conclusion implications and limitations of the study.

Chapter 2

Literature Review

2.1 Personalized pricing

While a wide variety of definitions of the term personalized pricing have been suggested, this study uses the term personalized pricing as a price discrimination approach that price is decided according to the customers' willingness to pay which is the characteristics of the consumer (OECD, 2018). An indispensable condition for this personalized pricing to take place is to employ data to measure consumer preferences (OFT, 2013; Bourreau, De Streel, & Graef, 2017). Companies need to have a system that collects and processes large amounts of data about consumers, which is associated with the high feasibility of personalized pricing in areas where data are easily gathered. In particular, in the case of online platforms, it is possible to collect data using their market power and use the collected data to increase market power again to create an entry barrier (Schepp & Wambach, 2016). Due to these characteristics, several attempts have been made for personalized pricing in e-commerce websites (Hannak, Soeller, Lazer, Mislove, & Wilson, 2014; Rafi, 2017).

Personalized pricing has received great attention because the concept of it is increasing revenue by finding consumers who can pay more than previous uniform prices and setting higher prices to them (Elmachtoub, Gupta, & Hamilton, 2018;

Shiller, 2014). In addition, this has an effect of expanding the market size by setting customized prices to consumers who want to pay less than the existing uniform price but are willing to pay more than the cost (OECD, 2016). However, personalized pricing has not only positive aspects but also negative aspects. Consumers may have a perception that the price given to them is unfair, which reduces the purchase probability and damage the brand image (Rayna, Darlington, & Striukova, 2015; Wagner & Eidenmuller, 2019; Gerlick & Liozu, 2020). In addition, consumers can have side effects such as acts of distorting data to prevent privacy infringement (Richards, Liaukonyte, & Streletskaya, 2016). Therefore, it is important for companies to forecast the ripple effect of the strategy by determining the advantages and disadvantages of personalized pricing strategy.

In this context, several studies have been conducted to implement personalized pricing effectively. They designed a model that is beneficial to both consumers and companies considering the advantages and disadvantages (Rayna et al., 2015), researched when to adopt personalized pricing strategy or uniform pricing strategy (Matsumura & Matsushima, 2015), examined the effects of consumer engagement in identity management (Chen et al., 2020), and investigated the strategies to be taken in consideration of data brokers (Montes et al., 2019).

However, there exist two major limitations in previous research. Firstly, most of the existing studies assume full observability of the reservation price of consumers when data are given. Considering the fact that it is impossible for a company to accurately determine the reservation price for all consumers and the company have to take actions based on the incorrectly predicted reservation price, it is necessary to focus on prediction performance. The study that estimates reservation price according to the characteristics of consumers and examines the effect of tailored price by using actual data is close to this study, but it also includes the above limitation (Shiller, 2014). Secondly, previous studies lack consumer interactions under personalized pricing strategies. One of the major concerns about personalized pricing is that consumers exchange information about their price and recognize their price given by the company is unfair. Therefore, consumers' interaction with price information is an important part. Accordingly, this study aimed to provide insights to companies about prediction performance by creating a personalized pricing model that takes consumer interactions into account

2.2 Agent-based models of innovation diffusion

Technology diffusion refers to the process by which the achievements of technological innovation are distributed to members through channels over time (Rogers, 2003). To model this diffusion process, aggregate models such as the Bass model have been used in the past, but it exists several problems. First, this method is not able to reflect the customer heterogeneity of individual preferences or behaviors. Second, it is difficult to reflect decision variables that enable what-if analysis which limits the use of the method. Third, it is impossible to consider interactions between consumers (H. Zhang & Vorobeychik, 2019; Ballot, Mandel, & Vignes, 2015). As a result, an agent-based modeling method that can observe emerging patterns based on individual decision-making processes has been risen.

Utility-based approach has been frequently used for modeling technology diffusion in the agent-based modeling technique. This method approaches that consumers make a decision by judging their utility for each alternative considering various factors. Agent-based models that reflect pricing strategy consider factors such as individual consumer characteristics, product quality, price, social impact, and promotion in their utility function (Delre, Jager, Bijmolt, & Janssen, 2007; Diao, Zhu, & Gao, 2011; Nejad, 2013; Karakaya, Badur, & Aytekin, 2011; Khouja, Hadzikadic, & Zaffar, 2008). In particular, in the case of personalized pricing, which is the subject of this study, the quality of products recognized by individuals can be replaced as a reservation price (Kowalska-Pyzalska, Ćwika, Jedrzejewski, & Sznajd-Weron, 2016).

In this study, an agent-based modeling technique is used to examine the diffusion process in the context of personalized pricing. The reason for using agent-based modeling is as follows. First, it is important to reflect the heterogeneity of consumers in the model because the personalized pricing strategy is based on individual preferences for products. Second, agent-based models can consider consumers' information exchanging behavior. Third, it is possible to change the value of the parameter and examine the impact in order to consider the effect of prediction performance. Detailed processes of constructing an agent-based model are covered in Chapter 3.

Chapter 3

Proposed procedure

3.1 Overview

In this chapter, we build an agent-based model from customer characteristics and purchasing behavior data, and examine how the personalized pricing strategy considering prediction performance affects the adoption behavior of consumers using the built model. We suggest a three-stage process as shown in figure 3.1.

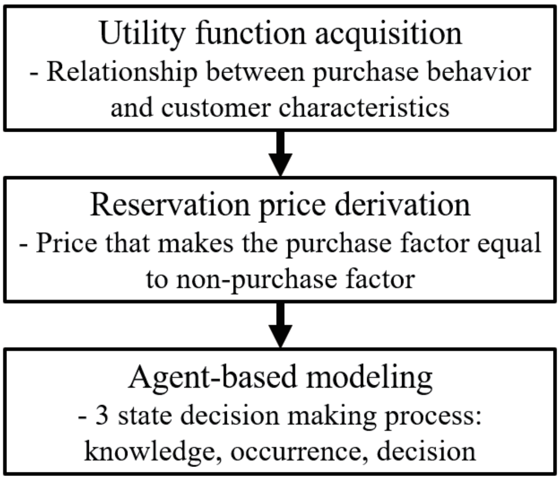

Figure 3.1: Overall process

The first step is to find a utility function that is the relationship between consumer characteristics and purchasing behavior. In this step, logistic regression analysis is performed using consumer characteristics as an input value and purchase or non-purchase as an output value. The first reason to use the logistic model is that the target variable in this model, adoption or non-adoption, is a binary value and input variable, input variables can be values with an infinite range. The second reason is that the target variable can be easily interpreted as a purchase probability in the logistic model and linear combination of the input variables can be interpreted as a utility

In the second step, the reservation price of the consumer is derived from the utility function obtained in the first step. This process is based on the fact that the purchase probability should be 0.5 when the purchase factor is the same as the non-purchase factor, and the utility is 0 in the logistic model. From this condition, it is possible to derive the price to make the utility zero, as the reservation price.

The final step is to construct an agent-based model in the context of personalized pricing. The constructed agent-based model reflects the process in which the consumer recognizes the product and calculates the utility to make a purchase decision when the purchase intention to the product arises. The factors affecting each step and the mathematical process of expressing them are described in Section 3.2.

3.2 Agent-based model

This agent-based model reflects the process in which a product is adopted by consumer agents when only a single product exists. The consumer decision-making process consists of the following steps: knowledge, occurrence and decision. This is in accordance with the existing consumer purchase decision model, where the consumer recognizes the need, acquires information and evaluates the utility value to

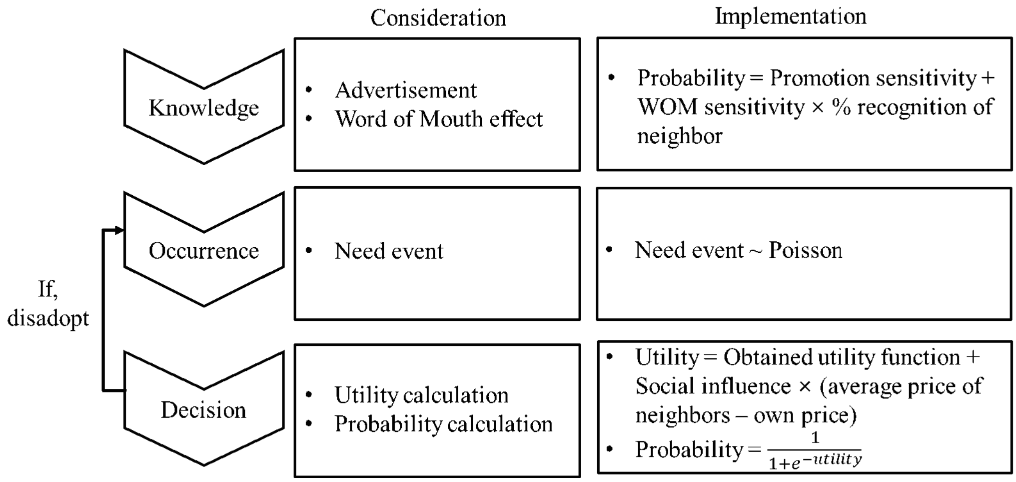

Figure 3.2: Decision-making process of individual customers

make a purchase decision (Rogers, 2003; T. Zhang & Zhang, 2007).

The first step is knowledge, the recognition of the existence of a product. There are two factors influencing this process: advertisement and word of mouth effect. The consumer recognizes the product through the company's marketing campaign or neighbors' word of mouth. Social sensitivity of each person is different depending on the heterogeneity of the individual consumer, so the parameter is set as a random variable that follows a uniform distribution [0,1].

The second step is the occurrence, the occurrence of need events. This step exists due to the fact that even if the consumer recognizes the product, in order to purchase the product, the consumer must have the need for the product. This process is modeled as using the Poisson distribution, which is a distribution that is frequently used when an existing event occurs.

The final step is the decision, the evaluation of the needed product. Consumers who recognize and demand the product calculate the utility of the product and determine the purchasing probability according to the calculated utility. The utility function is a linear combination of independent variables including price, and it is substituted into a logistic function to calculate the probability of purchase. Social influence, which is one of the important factors influencing consumer's purchasing decisions, is modified in consideration of personalized pricing situations and utility function can be derived from data on the web.

3.3 Prediction performance

This paper is based on the fact that it is difficult for companies to accurately predict the reservation price of consumers. In order to examine the effects of prediction performance, it is necessary to model how close the estimated reservation price is to the consumer's true reservation price. In this study, the concept of accuracy and precision, which are commonly used to analyze prediction performance, are introduced. As shown in Figure 3.3, accuracy is a criterion indicating how close the predicted value is to the actual value, and precision is a criterion indicating how close the results of multiple predicted values are to each other.

In response, two parameters, Mean and Variance, are introduced from the concept of accuracy and precision. Mean indicates the difference between the predicted value and the actual value and Variance indicates the variance of the predicted values. If a company uses a personalized pricing strategy, the price will be different for each consumer and the differentiated price will be applied to the price term of the utility function mentioned in the above step. Accordingly, in this model, prediction performance is expressed as follows:

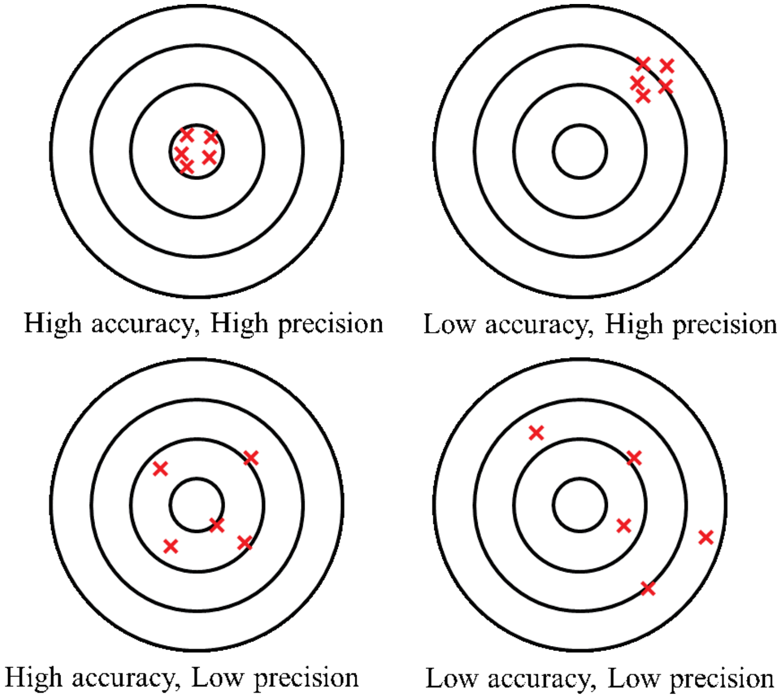

Figure 3.3: Concept of accuracy and precision

$$Price \sim Uniform[Reservation\ price\ +\ Mean - \frac{Variance}{2},$$
$$Reservation\ price\ +\ Mean\ +\ \frac{Variance}{2}]$$

For instance, if the reservation price of the consumer is 0, the Mean is 0.2, and the Variance is 0.2, the consumer's price is randomly decided from a Uniform distribution [0.1, 0.3]. The most accurate situation is when Mean is 0 and Variance is 0. At this time, the price that the company charges are the same as the reservation price of the consumer. If the Mean is far from 0 and the Variance is greater than 0, the set price is inconsistent with the reservation price of the consumer.

Chapter 4

Case study

A case study is conducted using e-commerce behavior data from a large multi-category online store published on Kaggle, the most popular data science competition platform. E-commerce is the field with the highest interest in personalized pricing and there are some practical application examples. Especially, given data is a form of e-commerce data set acquired by real companies, so it can be widely applied. This dataset is provided by the REES46 Marketing Platform and uses data from October 2019. Based on whether there is enough data to be analyzed, whether there is a difference in reservation price for each individual, and whether there exists multi-price data, Samsung Notebook with 43,071 rows was selected. The contents of the collected data are shown in the following table 4.1.

4.1 Utility function acquisition

Figure 4.1 presents the process of obtaining a utility function from an e-commerce dataset. The left half of the figure shows the process of converting the data into a form that can be inputted into the machine learning model, and the right half of the figure illustrates the process of assigning the processed input data to the Logistic

Table 4.1: File structure

Feature	Meaning
event time	Time when event happened at (in UTC)
event type	Events can be: view - a user viewed a product, cart - a user added a product to shopping cart, purchase - a user purchased a product
product id	ID of a product
category id	Product's category ID
category code	Product's category taxonomy (code name)
brand	Downcased string of brand name
price	Float price of a product
user id	Permanent user ID
user session	Temporary user's session ID

regression model to learn and evaluate the performance.

To begin this procedure, we perform preprocessing on the given data. It checks whether missing values exist, and appropriately fill or delete missing values. Also, dummy variables are introduced and categorical data converted to numerical types depending on the characteristics of variables because categorical data cannot be used as input values. Unlike most variables that have values between 0 and 1, the price variable has a value between 669 and 720. It indicates that there exists a possibility to learn the importance of the variable incorrectly and min-max scaling is applied to the price variable to prevent the problem.

Next, feature engineering is conducted. For product id, category id, category code, and brand variables, all rows in the dataset have the same values, so the corresponding columns are dropped. the user id and user session columns are deleted after their usage as an index disappears. Event time is an important column that includes useful information about customer behavior. New features are created to inform which time zone and which day of the week customer was visited by dividing

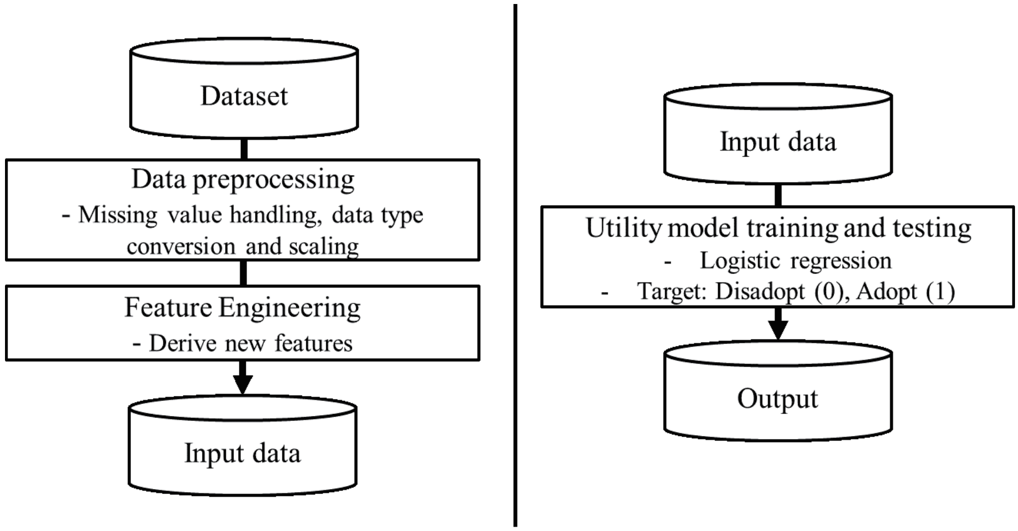

Figure 4.1: Utility function acquisition process

the column.

The processed data is converted so that each row represents each user's characteristics, and as a result, 19,119 rows are obtained. These data are unbalanced data with a probability of purchase of 3.049%, and oversampling is performed to adjust them. The calibrated data are separated by random sampling of train data: test data at a ratio of 80:20, and train data was inputted for the logistic regression model to train the model.

Table 4.2: Model evaluation result

Accuracy	**Precision**	**Recall**	**F1 Score**	**AUC**	**Confusion matrix**
0.9786	0.9892	0.9023	0.9437	0.9833	$\begin{bmatrix} 3669 & 9 \\ 89 & 822 \end{bmatrix}$

The evaluation results for the trained model are presented in table 4.2. The trained model predicts customers' adoption choice with 97.86% accuracy. From the confusion matrix, it is found that the false-negative value is higher than the false positive value. It means that the model tends to misjudge customers' decisions as non-adoption. This result is consistent with results where precision value is higher than the recall value.

$$
\begin{aligned}
utility = 5.42 - 10.05 \times Price\ scaled + 1.88 \times Event\ cart + 0.11 \times Event\ view \\
+ 0.25 \times Day\ 1 + 0.55 \times Day\ 2 + 0.14 \times Day\ 3 + 0.14 \times Day\ 4 \\
- 0.15 \times Mon - 0.47 \times Tue - 0.27 \times Wed + 0.11 \times Thu \\
- 0.36 \times Fri - 0.46 \times Sat - 0.49 \times Sun
\end{aligned}
$$

The result of the utility function from the trained model is as above. A total of 14 features are used as an input, and the importance of the features and the correlation can be known from the sign and size of the coefficients. The most remarkable feature is Price scaled, which indicates that it has a great influence on the utility because it has the largest coefficient value and negative correlation between price and utility is revealed from sign of the coefficient.

4.2 Reservation price derivation

Reservation price means the highest price that a buyer is willing to pay. If the price is the same as the reservation price, the adoption probability will be 0.5 due to the fact that motivation to adopt is the same as motivation to disadopt. The utility is 0 when the probability is 0.5, and the reservation price for each customer is driven by substituting this condition into the utility function obtained above. The reservation price result derived by applying the above utility model to each consumer's profile is as shown below, and it can be seen that the result follows a lognormal distribution. This is consistent with economic variables such as personal income or wealth follow lognormal distributions (Cantono & Silverberg, 2009). As a result of fitting the Price scaled variable to the lognormal distribution, is as follows: ln(Price scaled) $\sim N(-0.564, 0.126)$

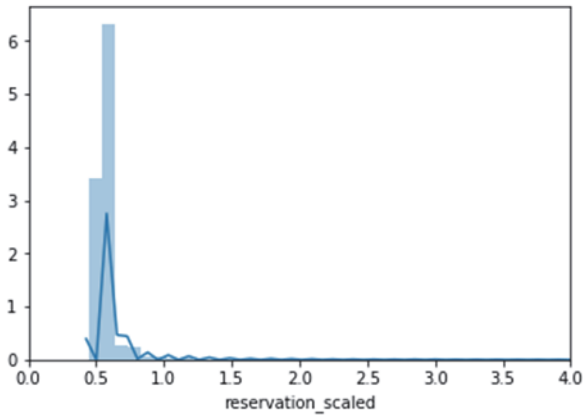

Figure 4.2: Reservation price result

4.3 Agent-based modeling

An agent-based model is constructed on the above results. Netlogo 6.0.4 software is employed for implementation and the parameters included in the model are summarized in Table 4.3. Based on these parameters, the consumer recognizes the product, evaluates the utility when a purchase need occurs, and makes a purchase decision.

Table 4.3: Agent-based model parameters

Parameter	**Meaning**	**Value**
Population	Number of agents in the model	1000
Network	Agent's social network type	Small world network
Need event	Need event of the agents	Poisson(0.005)
Strategy	Firm's pricing strategy type	[Uniform pricing, Personalized pricing]
Production Cost	Cost needed to produce the product	Based on the scenario, default 0.5
Reservation Mean	Difference between the mean of the estimated reservation price range and the actual reservation price	Based on the scenario, default 0
Reservation Variance	Variance of the estimated reservation price range	Based on the scenario, default: 0.2
Social Sensitivity	Sensitivity of the agents among neighbors information	If Positive: Uniform (0,1), If Negative: 3*Uniform (0,1)
WOM Sensitivity	Coefficient of the agents among neighbors word of mouth	Uniform(0, 0.02)
Promotion Sensitivity	Sensitivity of the agents about firm's promotion	Uniform(0, 0.1)

Environmental and agents' parameters are set before performing the simulation. For the environmental parameter, the total population is 1000, the simulation time is set to 300 days, and the need event follows Poisson (0.05). These values are set to trial and error. A small-world network is used to describe the agent's social network which is widely applied as the agent's social network (Watts, 2004). For the agent's parameter, social sensitivity, WOM coefficient, and promotion sensitivity exist. Their detailed values were also obtained through trials and errors. In particular, social sensitivity is specified in positive and negative impact cases to reflect the customers' perception. If the social sensitivity has a negative impact, the parameter value is tripled than positive case because consumers react more on loss than gains and it influences purchase probability more (Richards et al., 2016).

Chapter 5

Simulation and discussion

In this section, simulation is performed utilizing the constructed personalized pricing model. The purpose of this simulation is to understand the adoption of consumers according to each pricing strategy from the perspective of companies making inaccurate reservation price predictions.

To identify the impact of personalized strategy, we compared the uniform pricing strategy which sets the same price for all customers and the personalized pricing strategy which sets different prices for each consumer by predicting the reservation price. In the personalized pricing strategy, the company sets the price as the estimated reservation price when the estimated reservation price was higher than the production cost, and sets the price as the production cost when it is lower than the production cost. This is formulated as follows:

$$\text{Price} = \begin{cases} P_u,\ if\ a\ company\ is\ using\ uniform\ pricing \\ P_{ri},\ if\ a\ company\ is\ using\ personalized\ pricing\ and\ P_{ri}\ \geq\ P_p \\ P_p,\ if\ a\ company\ is\ using\ personalized\ pricing\ and\ P_{ri}\ <\ P_p \end{cases}$$

where,

$\mathrm{P}_u:\ price\ under\ uniform\ pricing\ strategy$

$P_{ri}:\ reservation\ price\ of\ consumer\ i\ under\ personalized\ pricing$

$P_p:\ production\ cost\ under\ personalized\ pricing$

Four scenarios are examined: the impact of production cost, social sensitivity, reservation price mean difference, and reservation price variance. The first scenario investigated the impact of production cost to identify net profit, not revenue. The second scenario considered the effect of word of mouth for the price, which is a major obstacle to the implementation of personalized pricing strategy. Third and fourth scenarios are both for examining prediction performance in perspective of how close the predicted reservation price is to the real value and how close the prediction values are to each other.

5.1 Impact of production cost

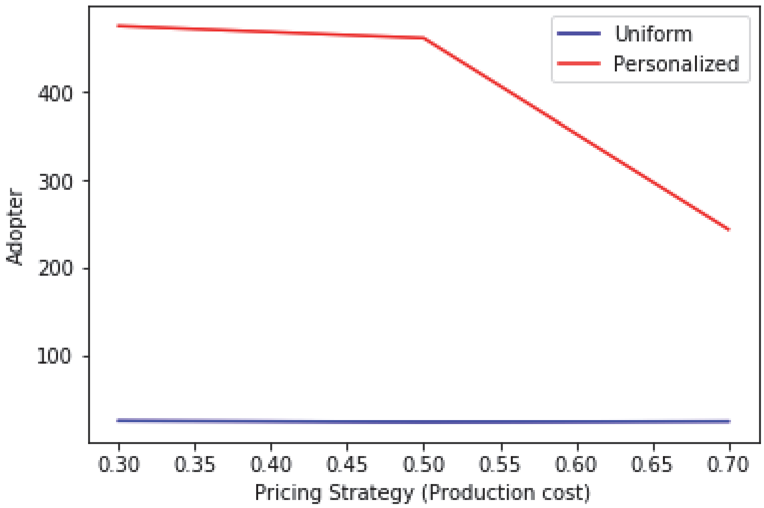

Figure 5.1: Number of adopters under each strategy

Figure 5.1 compares the number of adopters in the uniform pricing strategy and the personalized pricing strategy when the production cost changes from 0 to

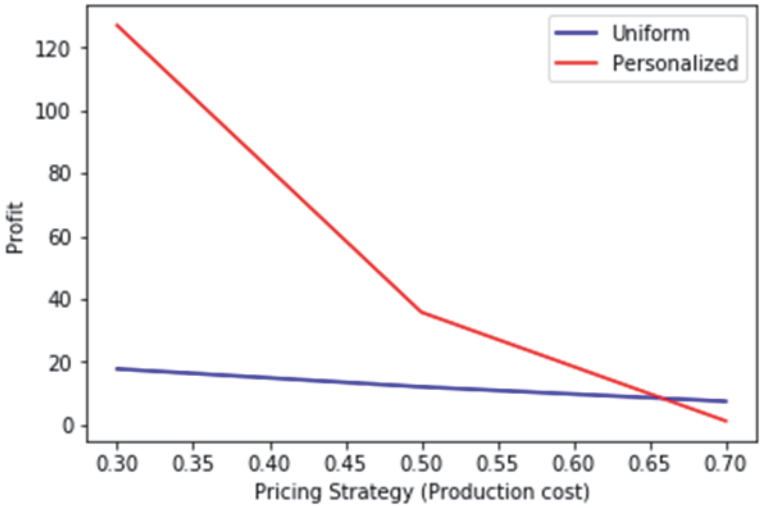

Figure 5.2: Net profit under each strategy

1. Figure 5.2 shows the change in profit in the same situation. The first interesting result is the correlation between production cost and the number of adopters and the correlation between production cost and profit. In the case of uniform pricing, there is no significant change in the number of adopters and profit when the production cost changes, but in the case of personalized pricing, it can be seen that the number and profile of the adaptors decrease rapidly as the production cost increases. In addition, personalized pricing case always shows a considerably larger number of adopters than the uniform pricing case. The reason is that the personalized pricing strategy can capture a broad range of customers. It is able to target customers who have reservation price lower than the uniform price, but higher than the production cost. However, in the perspective of profit, personalized pricing does not always benefit from uniform pricing. In Figure 5.2, it can be seen that the profit reduction due to the increase in production cost is greater in personalized pricing than in uniform pricing. This is because as production cost increases, customers with lower

reservation prices are priced at higher to the production cost, reducing the likelihood of their adoption.

5.2 Impact of social influence

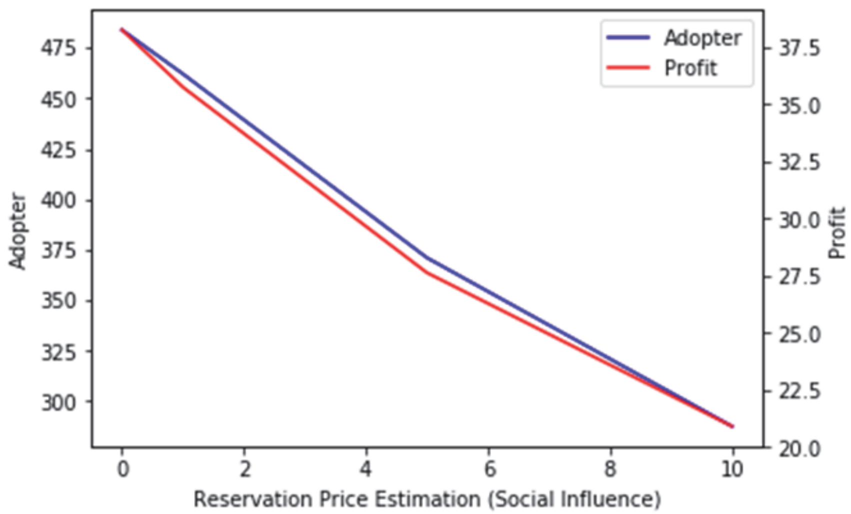

Figure 5.3: Number of adopters and net profit from social influence change

Figure 5.3 illustrates changes in the number of adopters and profits under the personalized pricing strategy as social sensitivity increases. From the previous utility function, social influence was described as a social sensitivity * (average price of neighbors – own price). This is based on the fact that, unlike the uniform pricing strategy, consumers can exchange information about the price they receive under the personalized pricing strategy. This part is modeled to have zero value in the uniform pricing strategy, positive value when the average price of neighbors is higher than their own in the personalized pricing strategy, and a negative value when the average price of neighbors is lower than their received price. As mentioned in Section 3.2, if the average price is lower than its own price, the weight of the coefficient increase because consumers respond more to negative word of mouth. Consequently, both the number of adopters and the net profit decrease as the social influence coefficient increases. This means that when using the personalized pricing strategy, companies should check whether consumers have a high or low social influence on price.

5.3 Impact of reservation price prediction performance

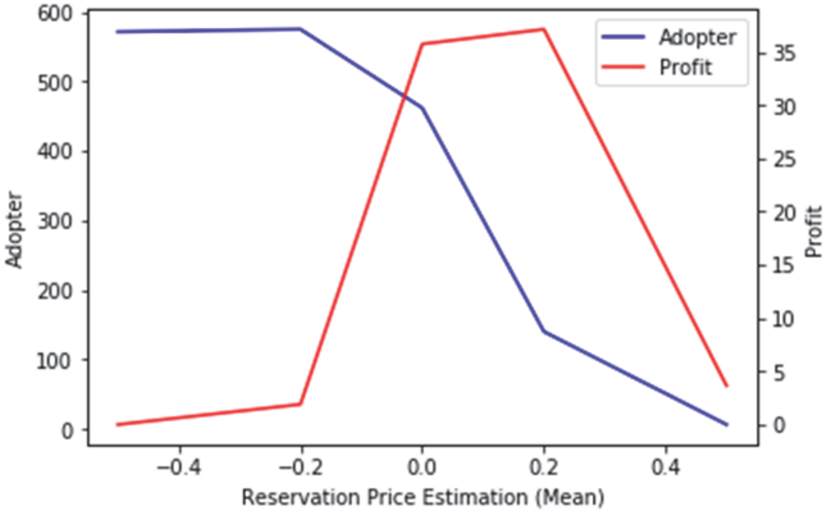

Figure 5.4: Number of adopters and net profit from mean change

Figure 5.4 presents the change in the number of adopters and profit according to the degree to which the average of the estimated reservation price deviates from the true reservation price. Figure 5.5 presents the change in the number of adopters and profit according to the variance of the estimated reservation price. As mentioned in Section 3.3, the most accurately predicted situation is a situation where the Mean is set as zero and the Variance is set as zero, which is an ideal situation

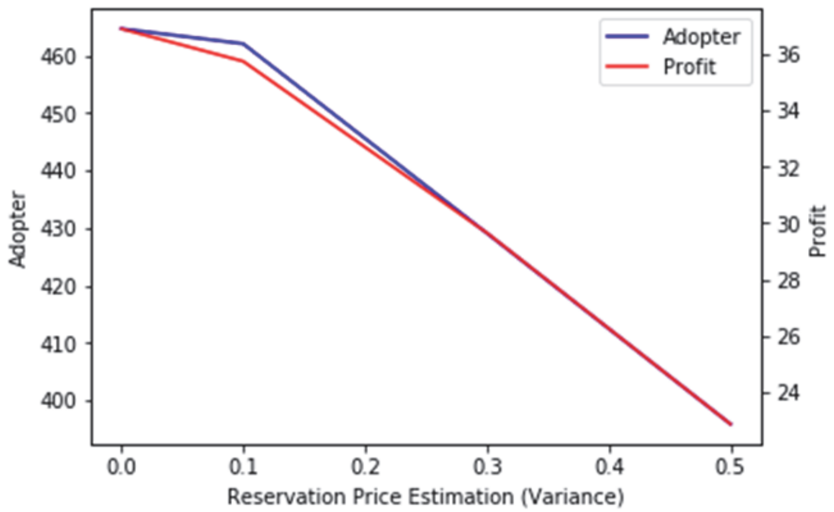

Figure 5.5: Number of adopters and net profit from variance change

where reservation price is perfectly predicted. The Mean variable can have a negative value because the average of the predicted values can be lower or higher than the reservation price, and the Variance variable always has a positive value because it means the variance of the predicted value.

From the data in Figure 5.4, it is apparent that the number of adaptors decreases as the mean increases from negative to positive. The reason is that a lower estimation of the reservation price leads to a higher adoption probability of consumers and a higher estimation of reservation price leads to lower adoption probability of consumers. Closer inspection of the profit shows the setting price as a reservation price might not be an optimal strategy for companies. It is seen that the profit is similar or rather slightly higher when the Mean is 0.2 than the Mean is 0 because even if the number of adopters decreases, the profit from each adopter increases. In addition, the value of the profit rapidly decreases in the range other than when the forecast mean is between 0 and 0.2, indicating that the application of the personalized pricing strategy may result in a lower profit if the accuracy of the mean is not guaranteed to some extent.

From the data in Figure 5.5, an increase in the variance of the predicted value is negatively correlated with the number of adopters and profits. Considering that Price is randomly charged from Uniform [reservation price + mean − variance/2, reservation price + mean + variance/2], the Mean have a greater impact on the adoption process when the values of variance and mean change. This means that the accuracy of the Mean variable is important than the accuracy of the Variance variable when predicting reservation price.

Chapter 6

Conclusion

The aim of the present research is to propose a personalized pricing agent-based model considering prediction performance that was not considered in the existing research. It is impossible to perfectly predict customer's reservation price and reservation price prediction performance is directly related to the company's profits, it is important to determine the effects of prediction performance. Also, the main obstacle for implementing personalized pricing strategy is the customer's perception which is formed by social interaction. Therefore, this study modeled the decision-making behavior of consumers under a personalized price strategy in consideration of social influence on prices and prediction performance. Then, an agent-based model was constructed using e-commerce data. Finally, simulations are conducted to identify the impact of the pricing strategy, and the results are discussed.

As a result of the case study, a personalized pricing strategy is not always superior to a uniform pricing strategy. Investigation on production cost reveals that personalized pricing strategy has the effect of expanding the market compared to a uniform pricing strategy, but it does not always result in more profit. Analysis of the social influence shows that social influence about price has a negative effect on actual adoption, companies need to analyze consumers' social characteristics. In the case of prediction performance, the mean difference is an important factor that significantly affects the number of adopters and profit, and when it is slightly higher than the actual reservation price, the number of adopters decreases, but profit increases slightly. In addition, the mean difference value more than a certain degree results in a rapid decrease in profit. In the case of the variance of the predicted values, the variance is negatively correlated to both the number of adopters and the profit. These results indicate that it is meaningful to apply a personalized pricing strategy when the predicted performance is above a certain level.

This study exists several limitations. First, this study assumes the monopoly market situation, but a notebook case, a competitive market product, is selected and analyzed. The selection reason is that although most of the products sold in e-commerce are products in the competitive market, e-commerce is still a platform to acquire and utilize large amounts of data about consumers. Second, the validation issue of research results exists. In this study, the reservation price is derived using the data and a personalized model simulation is conducted based on the result. Derived reservation price distribution shows a lognormal distribution like the existing economic variable, but simulation results are not verified with the actual personalized pricing results. In many cases, companies try to hide the fact that they are using personalized pricing strategy to customers because of consumer resistance. Therefore, it is hard to collect real data for validation. Accordingly, the proposed agent-based model is established based on the existing research results to reflect the validation issue.

References

Ballot, G., Mandel, A., & Vignes, A. (2015). Agent-based modeling and economic theory: where do we stand? *Journal of Economic Interaction and Coordination*, *10*(2), 199–220.

Bourreau, M., De Streel, A., & Graef, I. (2017). *Big data and competition policy: Market power, personalised pricing and advertising* (Project Report). CERRE.

Cantono, S., & Silverberg, G. (2009). A percolation model of eco-innovation diffusion: the relationship between diffusion, learning economies and subsidies. *Technological forecasting and social change*, *76*(4), 487–496.

Chen, Z., Choe, C., & Matsushima, N. (2020). Competitive personalized pricing. *Management Science*.

Delre, S. A., Jager, W., Bijmolt, T. H., & Janssen, M. A. (2007). Targeting and timing promotional activities: An agent-based model for the takeoff of new products. *Journal of business research*, *60*(8), 826–835.

Diao, J., Zhu, K., & Gao, Y. (2011). Agent-based simulation of durables dynamic pricing. *Systems Engineering Procedia*, *2*, 205–212.

Elmachtoub, A. N., Gupta, V., & Hamilton, M. (2018). The value of personalized pricing. *Available at SSRN 3127719*.

Gerlick, J. A., & Liozu, S. M. (2020). Ethical and legal considerations of artificial intelligence and algorithmic decision-making in personalized pricing. *Journal of Revenue and Pricing Management*, 1–14.

Hannak, A., Soeller, G., Lazer, D., Mislove, A., & Wilson, C. (2014). Measuring price discrimination and steering on e-commerce web sites. In *Proceedings of the 2014 conference on internet measurement conference* (pp. 305–318).

Karakaya, Ç., Badur, B., & Aytekin, C. (2011). Analyzing the effectiveness of marketing strategies in the presence of word of mouth: agent-based modeling approach. *Journal of marketing research and case studies*, *2011*, 1–17.

Khouja, M., Hadzikadic, M., & Zaffar, M. A. (2008). An agent based modeling approach for determining optimal price-rebate schemes. *Simulation Modelling Practice and Theory*, *16*(1), 111–126.

Koen, V. (2014, Sep). Getting to know you. *Economist*. Retrieved from `https://www.economist.com/special-report/2014/09/11/getting-to-know-you`

Kowalska-Pyzalska, A., Ćwika, K., Jedrzejewski, A., & Sznajd-Weron, K. (2016). Linking consumer opinions with reservation prices in an agent-based model of innovation diffusion. *Acta Physica Polonica, A.*, *129*(5).

Martin, N. (2019, May). Uber charges more if they think you're willing to pay more. *Forbes*. Retrieved from `https://www.forbes.com/sites/nicolemartin1/2019/03/30/uber-charges-more-if-they-think-youre-willing-to-pay-more`

Matsumura, T., & Matsushima, N. (2015). Should firms employ personalized pricing? *Journal of Economics & Management Strategy*, *24*(4), 887–903.

Montes, R., Sand-Zantman, W., & Valletti, T. (2019). The value of personal information in online markets with endogenous privacy. *Management Science*, *65*(3), 1342–1362.

Nejad, M. G. (2013). Optimal pricing for the growth of innovations with direct network externalities: an agent-based approach. *Journal of Product & Brand Management*.

OECD. (2016). Price discrimination. Retrieved from https://one.oecd.org/document/DAF/COMP(2016)15/en/pdf

OECD. (2018). Personalised pricing in the digital era. Retrieved from https://one.oecd.org/document/DAF/COMP(2018)13/en/pdf

OFT. (2013). Personalised pricing – increasing transparency to improve trust. Retrieved from https://webarchive.nationalarchives.gov.uk/20140402165101/http://oft.gov.uk/shared_oft/markets-work/personalised-pricing/oft1489.pdf

Rafi, M. (2017). How retailers use personalized prices to test what you're willing to pay. *Harvard Business review*, *20*.

Rayna, T., Darlington, J., & Striukova, L. (2015). Pricing music using personal data: mutually advantageous first-degree price discrimination. *Electronic Markets*, *25*(2), 139–154.

Richards, T. J., Liaukonyte, J., & Streletskaya, N. A. (2016). Personalized pricing and price fairness. *International Journal of Industrial Organization*, *44*, 138–153.

Rogers, E. M. (2003). *Diffusion of innovations*. Free Press.

Schepp, N.-P., & Wambach, A. (2016). On big data and its relevance for market power assessment. *Journal of European Competition Law & Practice*, *7*(2), 120–124.

Shiller, B. R. (2014). *First degree price discrimination using big data* (Working Paper). Waltham, MA: Department of Economics, Brandeis University.

Wagner, G., & Eidenmuller, H. (2019). Down by algorithms: Siphoning rents, exploiting biases, and shaping preferences: Regulating the dark side of personalized transactions. *University of Chicago Law Review*, *86*, 581.

Watts, D. J. (2004). *Small worlds: the dynamics of networks between order and randomness* (Vol. 9). Princeton university press.

Zhang, H., & Vorobeychik, Y. (2019). Empirically grounded agent-based models of innovation diffusion: a critical review. *Artificial Intelligence Review*, 1–35.

Zhang, T., & Zhang, D. (2007). Agent-based simulation of consumer purchase decision-making and the decoy effect. *Journal of business research*, *60*(8), 912–922.

국문초록

디지털 환경의 발달로 인해 개인화된 가격 전략 (personalized pricing)이 많은 관심을 받고 있다. 과거와는 달리 오늘날에는 데이터 수집 환경, 분석 알고리즘과 하드웨어 발달로 개별 소비자들의 선호를 점점 파악하기 용이해지면서 소비자의 유보 가격 (reservation price)을 예측해서 가격을 매기는 개인화 된 가격 전략이 실행 가능한 환경이 되었다. 기업들은 이 전략을 이용해서 시장을 넓히고 더 많은 수익을 얻고자 이 주제에 많은 관심을 가지고 있다. 그러나 유보 가격을 정확히 예측하는 것이 불가능하고, 부정확한 유보 가격 예측으로 인해 소비자들에게 잘못된 가격을 제시하게 된다면 결국 역효과를 발생시킬 수 있다.

이에 따라 본 연구에서는 행위자 기반 모형 (agent-based model)을 세워 유보 가격 예측 정확성에 따른 개인화된 가격 전략에 따른 효과를 살펴본다. 먼저, 가격과 소비자의 구매 여부, 행동 데이터가 포함된 웹 데이터를 수집하여 가격을 포함한 소비자의 효용 함수를 구축한다. 다음으로 구축된 효용 함수를 이용해 소비자의 유보 가격을 도출한다. 그리고 이를 개별 소비자들의 의사결정 과정으로 하는 행위자 기반 모형을 만들어 시뮬레이션을 해본다. 행위자 기반 모형은 예측 정확성에 따른 효과를 조사할 수 있으며 개인화된 가격 전략 아래에서 소비자의 이질성과 상호작용을 고려할 수 있기 때문이다.

본 연구는 개인화된 가격 전략 아래에서 제품 확산에 대한 몇가지 중요한 통찰력을 제공한다. 첫째로, 이는 부정확한 유보 가격 예측에 따른 효과를 분석하여 기업들이 수익을 최대화하기 위한 통찰력을 제공한다. 둘째로, 기존의 단일화된 가격 전략 (uniform pricing)에서는 고려하지 않았지만 개인화된 가격 전력 상황에서는 고려해야하는 개별 소비자 가격에 대한 사회적 영향력을 포함한 행위자 기반 모형을 구축하여 기존의 확산

연구에 기여한다. 셋째로, 이는 실제 웹 상에서 발생되는 데이터를 활용하여 데이터 전처리, 효용 함수 도출, 모델 설립 및 전략 분석까지 전체 프로세스를 보여준다.

주요어: 개인화된 가격 전략 (Personalized pricing), 유보 가격 (Reservation price), 예측 성능, 확산, 행위자 기반 모형 (Agent-based model)

학번: 2017-28225