

Incorporating Cannibalization into Pricing Optimization Using Choice Data: An Application to the Pricing of Mobile Broadband Services

Fransiscus Ryan Pratikto^{1*}

Abstract: Price differentiation may not be as effective in increasing profitability due to imperfect segmentation, arbitrage, and, cannibalization. Cannibalization takes place when a customer with higher willingness-to-pay buys a lower-priced product. This research proposes an approach to incorporating cannibalization into pricing optimization using choice data. From choice data, individual-level utilities are estimated using hierarchical Bayes, and individual choice is predicted using randomized first choice simulation. Individual choices are then aggregated to obtain the demand function. The novelty of this research is in the way cannibalization is incorporated into the pricing optimization. Instead of integrating cannibalization into the demand function or representing it as a separate component in the optimization formulation, in this research, cannibalizing products are incorporated into the simulation scenario as competing products, based on which the demand functions used in the optimization are derived. This approach is more direct and realistic than those in previous research. The approach was implemented in a case study of mobile broadband services in Indonesian price-sensitive market. The result shows that two-fare-class price differentiation incorporated with product differentiation increases the total contribution of about 60% compared to the single-fare-class policy. Furthermore, it is also shown from our case study that starting from a three-fare-class policy, through iterations, our approach suggests that policy with two-fare-class results in a not significantly different total contribution.

Keywords: Cannibalization; pricing optimization; choice data; hierarchical Bayes; randomized first choice.

Introduction

Despite its central position among revenue management tactics, in certain conditions, price differentiation may not be as effective in increasing profitability. This may be due to imperfect segmentation, arbitrage, and, cannibalization (Philips [1]). Imperfect segmentation happens when the demand model cannot explain how customers make choices, while arbitrage occurs when price difference creates an opportunity for an arbitrageur to buy the product at a lower price and sell it at a higher price to customers with greater willingness-to-pay. Cannibalization takes place when a customer with higher willingness-to-pay buys lower price product. In general, arbitrage may be prevented by product attribute differentiation or non-attribute barriers, which is relatively easier to do. On the other hand, resolving the problem of imperfect segmentation and cannibalization required improvement in the way demand is modeled and estimated, which could be very challenging depending on the complexity of the product and market.

As an old adagium in management says “you cannot manage what you cannot measure”, measuring or estimating cannibalization is the key to resolving it. Lomax [2] proposed three alternative ways for measuring cannibalization, i.e. gains loss analysis, cross-purchase tables, and deviations from the expected share movements, all can be estimated from sales data. Those measures, while showing the magnitude of cannibalization, do not indicate the way to resolve it. What we need is a measure which relates cannibalization to controllable variables that account for it. Specifically, we need a demand model defined as a function of product attributes (including price) and other relevant variables (i.e. market segmentation variables). Cannibalization effect in a certain pricing scenario can be estimated once the demand function for the corresponding scenario is known.

Many researches in the past worked on developing demand functions which are capable of measuring cannibalization effect. Carpenter and Hanssens [3] developed an air travel passenger demand model in which the share of a fare-class is represented as a function of prices of all fare-classes using an exponential function. The fare-class share then goes into the price optimization problem, based on which the cannibalization effect can be estimated accordingly. Meredith and Maki [4] estimating cannibalization

¹ Faculty of Industrial Technology, Department of Industrial Engineering, Parahyangan Catholic University, Jl. Ciumbuleuit No. 94, Hegarmanah, Kec. Cidadap, Bandung, Jawa Barat 40141.
Email: frianp@unpar.ac.id

* Corresponding author

using demand model expressed as the natural exponential of a linear function of prices of competing products and difference between those prices. Srinivasan *et al.* [5] developed a model to forecast demand (i.e. Autoregressive Integrated Moving Average or ARIMA) which incorporates cannibalization. In this model, cannibalization is expressed as a function of product attributes and estimated from historical data. Albuquerque and Bronnenberg [6] developed a demand model capable of estimating cannibalization using aggregate logit. The model is parameterized using aggregate data (penetration and purchase data). Heerde *et al.* [7] proposed a demand model which incorporates cannibalization using time-varying vector error correction method. The model considers five sources of demand: (i) primary demand, (ii) cannibalization within the category, (iii) cannibalization between categories, (iv) brand switching within the category, and (v) brand switching between categories. Pacras *et al.* [8] developed a demand model for retail chain expansion which considers cannibalization between stores. In this model, demand is represented as a function of chain's goodwill (expressed as a function of current goodwill and advertising), occasions (e.g. holidays), store-level time-varying characteristics (e.g. price, satisfaction score), and store store's fixed effect (e.g. free-standing units (FSUs) versus mall, proximity to highways). The choice probability is estimated using logistic distribution. Haynes *et al.* [9] employed nested logit in a demand model, out of which elasticities can be calculated and used to estimate the cannibalization effect. The demand is a function of product attributes (including price) and share of a group where the product belongs. Ngwe [10] investigated cannibalization between regular and outlet stores using nested logit demand model, which is expressed as a function of product quality, product newness, price, type of outlet, time of purchase, and type of product. Guo and Chen [11] developed a demand model for a multi-generation product which considers cannibalization. In this research, demand is represented as the adoption rate, which increases linearly with the number of existing adopters. The model is expressed as a continuous-time differential equation. The model seeks to investigate the effect of a price discount on the demand of the multi-generation product. In Reimers and Xie [12], the effects of coupons on revenue cannibalization are investigated, where the number of cannibalization is assumed to be proportional to the difference between regular price and coupon price and to the inverse of coupon cost or disutility.

In all those researches, demand (or choice probability) is defined as a function of price, product attributes, and/or other segmentation-related variables. For product in a competitive setting, it is

plausible to explicitly include variables that represent competitors as in Carpenter and Hanssens [13], Meredith and Maki [4], and Pacras *et al.* [8]. To parameterize the model, all the above research use revealed preference data, both in aggregate and disaggregate terms. The advantage of using revealed preference data is its validity since it comes from actual or observed behavior. However, since revealed preference method evaluates current condition, it cannot be used directly to analyze situations that do not yet exist, which is a desirable property in a application such as pricing, marketing, and product development. On the other hand, while relatively more biased, stated preference data could be used to analyze hypothetical situations which could be very useful in such applications.

Basically, demand is the aggregation of customer's choices. In economic demand theory, random utility model (RUM) is considered the best approach to modeling choices [5]. Under this model, a decision-maker is assumed to be a utility-maximizer over several discrete alternatives, each of which has random utility (McFadden [14]). Some of the above research adopt this approach, such as Albuquerque and Bronnenberg [6], Pacras *et al.* [8], Haynes *et al.* [9], and Ngwe [10]. Modeling demand this way ensures its validity. However, when it comes to the optimization problem as in pricing, it results in a mathematically complex formulation which is difficult to solve analytically.

This research proposes an approach to incorporating cannibalization into pricing optimization problem using stated-preference data, specifically choice data. Compared to other types of stated preference data (i.e. ranking, rating), choice data has an advantage of better mimicking customer decision making a process in the market. Using choice data, individual-level utilities are estimated using hierarchical Bayes (Allenby *et al.* [15], which can identify the heterogeneity of respondents (Allenby and Ginter[16]). This method assumes multivariate normal utility values whose parameters are estimated iteratively using a method involving the Metropolis-Hastings algorithm. Having had individual level utility values, the choice from each respondent is simulated using randomized first choice method (RFC) [17]. RFC is based on RUM where two random components representing product variability and attribute variability are added. Individual choices are then aggregated to obtain a share of preference. The demand function is obtained by calibrating the share of preference with actual sales data. Cannibalization is incorporated into the demand function by including the cannibalizing products in the simulation scenario, based on which the demand function is derived. Using this demand function, optimal pricing

is determined accordingly. Most of the time, the optimal prices are different with price levels in the choice data. Iterations are needed to resolve this.

As in the previous research, attributes (including price) of all products in the market (both cannibalizing and competitor products) are considered. In this research, they are incorporated into the competitive scenarios based on which the simulation is run to obtain the demand function. The novelty of this research is in the way cannibalization is incorporated into the process of deriving demand function. Instead of integrating cannibalization into the demand function or representing it as a separate component in the formulation of an optimization problem, cannibalizing products are incorporated into the simulation scenario as competing products. Intuitively, this approach is more direct and realistic than those in the previous research.

We implement the approach in a case study of pricing optimization for mobile broadband services. It is not uncommon that a mobile broadband service provider offers more than one product for a market. Typical segmentation in this industry usually divides the market into clusters representing those that are price-sensitive, value-seeker, and premium. Some service providers have even more niched products. The segmentation is relatively unfenced such that those who are premium customers may end up buying products devoted to the price-sensitive market. Pricing optimization for mobile broadband service can be considered as an uncapacitated problem. The 4G, 4.5G, and 5G technology have made service providers can increase capacity significantly while keeping capital expenditure flat or even lower.

This paper is organized into four sections. This first section discusses the previous research and the novelty of this research, while the second one describes the methods. The third section elaborates the proposed approach into a case study and discusses the results, and the last part concludes.

Methods

This section discusses in an elaborate yet concise way, three topics that are fundamental to our problem, i.e. estimating individual utilities, market simulation and the demand function, and pricing optimization with cannibalization. Figure 1 depicts the sequential process of how choice data are used for estimating demand function and eventually, the pricing optimization.

In the first step, individual-level utilities are estimated from choice data using hierarchical Bayes me-

thod. In the next three steps, demand function is derived using the sequential processes of estimating individual choices, estimating share-of-preferences, and obtaining explicit and continuous demand function using interpolation. In the last step, the demand function goes into pricing optimization.

Estimating Utilities from Choice Data

Under RUM, a decision-maker will choose an alternative with the highest total utility. Once we know a decision maker's utility system, we can predict her preference towards any particular choice situation. Of course, in this case, the choice situations are limited to those that can be represented using the attributes and levels based on which the utility values are estimated. In our research, utilities are estimated from choice data. In this regard, choice data are responses to sets of balanced and orthogonal randomly generated choice tasks, each of which consists of a fixed number of stimuli. Each choice task may contain a "none" option. The questionnaires are designed and generated such that there are no two respondents receive identical sets of choice task. One or two fixed choice tasks are usually added in each set to test the internal validity of the utility model.

A common way to estimate utilities from choice data is using maximization of some likelihood function. This results in the parameters of the distribution of utility values in the aggregate term. Using aggregate utility values is like using mean and variance to represent the population. This approach can be extended to one with a number of clusters where clusters are represented with different means and variances. Another method involves estimating utility values for each individual. Having utility values for each individual is always better than just the average value of the population. This disaggregate approach usually assumes that utility values for each individual follow a multivariate normal probability distribution. The estimation is conducted through an iterative calculation using Markov chain Monte Carlo (MCMC) method. Hierarchical Bayes (HB) is considered the best method that falls into this category (Train [18]).

In HB there are two levels of estimation, i.e. estimation of individual utilities at the upper level, and, at the lower level, estimation of the probability that an individual chooses a particular alternative [19]. The vector representing utility values of individual j , β_j follow a multivariate normal distribution, $\beta_j \sim N(\mu, \Sigma)$, where μ is the vector of the means and Σ is the covariance matrix. In HB, β (aggregation of β_j for all individuals), μ , and Σ are estimated using the following algorithm:

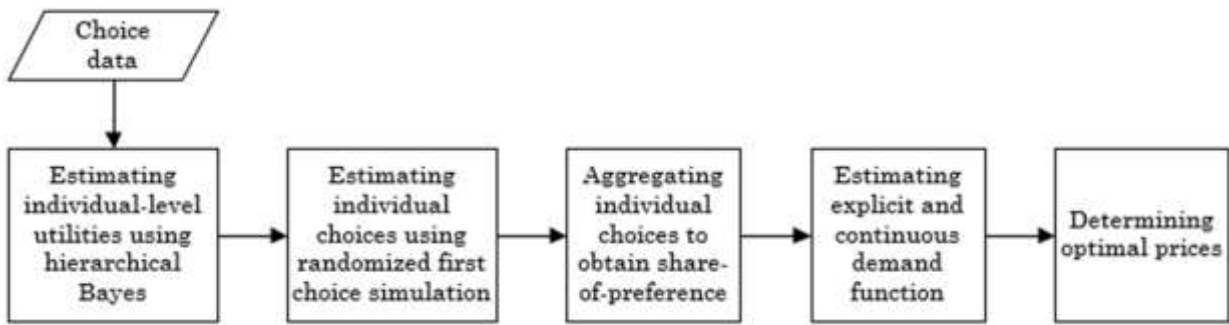


Figure 1. Pricing optimization with demand function derived from choice data

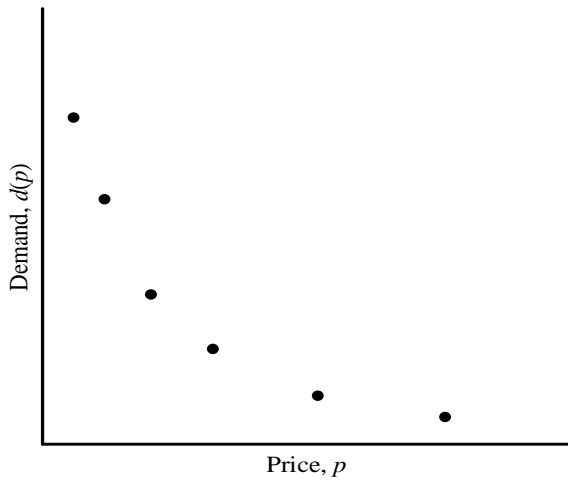


Figure 2. A typical plot of demand data points using hypothetical data

- Step 0 Set $t = 0$. Set initial value for β_t , μ_t , and Σ_t
- Step 1 Estimate μ_{t+1} by drawing from $N(\beta_t, \Sigma_t)$
1. Estimate Σ_{t+1} based on β_t and μ_{t+1} using Cholesky decomposition on the inverse of matrix $\mathbf{H} = \rho \mathbf{I} + \sum_n (\mu_{t+1} - \beta_t)(\mu_{t+1} - \beta_t)^{-1}$
 2. Estimate β_{t+1} based on μ_{t+1} and Σ_{t+1} using Metropolis-Hastings algorithm
- Step 2 If β , μ , and Σ have converged, stop. Otherwise, set β_{t+1} , μ_{t+1} , and Σ_{t+1} as β_t , μ_t , and Σ_t , respectively, and go to Step 1.

Market Simulation and the Demand Function

Up to this point, we already have vector β_j which represents utilities of all attribute levels for individual j . Using this information, we can predict how individual j would choose in a particular choice situation. The attributes and levels are predetermined during the study design and kept fixed during the estimation and market simulation.

Given the utility value of each attribute level of each decision-maker, a number of methods can be used to predict how she would choose in a particular choice situation. They are the first choice, logit model, dan randomized first choice. Randomized first choice or RFC is considered the best method for predicting

how an individual makes choices given her utility values for all attribute levels [20]. This is because it is immune from the independence-of-irrelevant-alternatives problem, and more realistic since it incorporates randomness in modeling how people choose.

Let \mathbf{X}_i be the binary vector of attribute level of product i , and U_{ij} be the utility of product i for individual j . According to RFC.

$$U_{ij} = \mathbf{X}_i(\beta_j + \mathbf{E}_p) + E_f \quad (1)$$

where \mathbf{E}_p is a vector representing product variability, while E_f is a scalar representing individual variability, both follow Gumbel probability distribution. \mathbf{E}_p and E_f are determined such that the simulator can accommodate differential impact, differential substitution, and differential enhancement [10]. Having U_{ij} simulated based on the above equation, individual j will choose product i in a particular choice scenario if $U_{ij} > U_{ik}, \forall k \neq j$. If these choices are aggregated for all respondents, we can estimate the choice proportion of each product in the chosen scenario. Let π_i be the share-of-preference, i.e. the proportion of respondents that choose product i in a choice scenario represented as set Ω . We refer Ω as the concepts other than our concept of interest. This means that the simulation scenario is the union of Ω and our concept of interest (ω), expressed as $\{\Omega, \omega\}$. For simplicity, we define RFC_i as a function that converts information about utilities of each individual and attribute levels of product i to proportion of respondents that choose product i in choice scenario Ω , as described above. If β (aggregation of β_j for all j) is the matrix representing utilities of all respondents, then we have

$$\pi_i = RFC_i(\mathbf{X}_i, \beta | \{\Omega, \omega\}), \forall i \in \{\Omega, \omega\} \quad (2)$$

Suppose there are m levels of price (p_1, \dots, p_m) based on which β is defined, and let $\mathbf{X}_i = [\mathbf{X}_i^{(o)}, X_i^{(p_1)}, \dots, X_i^{(p_m)}]$ where $\mathbf{X}_i^{(o)}$ be a vector that represents the elements of \mathbf{X}_i that corresponds to

attributes other than price, and $X_i^{(p_1)}, \dots, X_i^{(p_m)}$ be the elements that correspond to price levels p_1, \dots, p_m , respectively. Now let $\pi_i^{(p_l)}$ be the proportion of respondents that choose product i when it is offered at price of p_l . We can estimate $\pi_i^{(p_l)}$ by setting $X_i^{(p_l)}$ to 1, and 0 for other price levels, running the RFC simulator and calculating the corresponding share of preference.

$$\pi_i^{(p_l)} = RFC_i \left(\left[\mathbf{X}_i^{(p_o)}, X_i^{(p_l)} = 1, \mathbf{X}_i^{(p_o)} = 0 \right], \boldsymbol{\beta} \mid \{\Omega, \omega\} \right) \quad (3)$$

where $\mathbf{X}_i^{(p_o)}$ represents vector of elements that correspond to price levels other than l . If we run the simulation for each price levels, we will have m data points

$$(p_1, \pi_i^{(p_1)}), (p_2, \pi_i^{(p_2)}), \dots, (p_m, \pi_i^{(p_m)}).$$

From these share of preference data points, we can get demand data points by multiplying the share of preferences with a quantity representing the achievable market size, D . If we make it more general by removing the index i and letting demand at the price of p_k , $d(p_k) = \pi^{(p_k)} \cdot D$ for $k = 1, \dots, m$, then we will have m demand data points $(p_1, d(p_1)), (p_2, d(p_2)), \dots, (p_m, d(p_m))$. In pricing optimization, we use the inverse demand curve to represent demand function because the price is the decision variable. If the choice data and the estimation process are good, and our product of interest is a normal product, we will come up with a plot of demand data points that looks like one in Figure 2.

To make these data points ready for use in the pricing optimization, we need to convert them into an explicit demand function. This involves interpolating and/or fitting the data points to a particular function. Under deterministic assumption we can do this in two ways: (1) fitting the data points to one assumed demand function, and (2) interpolating between two consecutive data points using splines of a certain function. For the first approach, readers may refer to Huang *et al.* [21] for alternative demand functions that can be used in the fitting. This approach results in a realistic demand function but may not fit with all data points. The second approach will produce a function with the perfect fit, but the function may be too complex for optimization. One of the functions for the splines that will produce a realistic and perfect fit demand function is cubic [22].

Pricing Optimization

Basically, pricing optimization seeks to determine price levels that maximize total contribution, subject to constraints such as capacity or other restrictions. The basic formulation of uncapacitated single-segment pricing optimization is as follows [1].

$$\begin{aligned} \max_p \quad & d(p)(p - c) \\ \text{s.t.} \quad & p \geq 0 \end{aligned} \quad (4)$$

where p is price, $d(p)$ is demand as a function of price, and c is the incremental cost. The complexity of this problem depends on the nature of the demand function. For two-segment pricing, the formulation becomes [23].

$$\begin{aligned} \max_{(p_1, p_2, v)} \quad & m(p_1, p_2; v) = \min(d(p_1), d(v))(p_1 - c_1) + \max(d(p_2) - d(v), 0)(p_2 - c_2) \\ \text{s.t.} \quad & 0 < p_2 \leq v < p_1 \end{aligned} \quad (5)$$

where p_1 and p_2 are prices for higher-WTP and lower-WTP segment, respectively, while v is the WTP boundary between the two segments. The incremental costs are expressed as c_1 and c_2 to allow for product differentiation. To be effective, price differentiation is often delivered along with product differentiation. The formulation in equation 5 assumes that demand from those segments is independent. This means that consumers are perfectly segregated based on their WTP and those having higher WTP could not access the product sold at a lower price. This optimization problem uses one demand function which can be derived using RFC simulation (as described in the previous section) under a particular scenario. In this case, the scenario is scenario $\{\Omega, \omega\}$ in which ω corresponds to our concept of interest with the demand function of $d(p)$. This single demand function is then divided into two sections, one covers the higher price region (p_1 region), and the other covers the lower region (p_2 region).

In reality, this independent demand assumption often does not hold. Although sellers may use some restrictions to impose price differentiation tactic (based on ages, time, place, etc.), in many circumstances, consumers with higher WTP still can get access to the lower-priced product (cannibalization). When this happens, demand for a lower-price product will increase while that for higher-price will decrease. The magnitude of the change depends on the disparity between the two prices. Hence, we cannot assume that those two segments are independent because, based on the method we use to derive the demand function, demand from consumers with higher-WTP depends on the product price for a lower-WTP segment, and vice versa. This means that, instead of one, now we have two demand functions, $d_1(p_1)$ for higher-WTP consumers and $d_2(p_2)$ for lower-WTP consumers. The corresponding choice scenario is now $\{\Omega, \omega_1, \omega_2\}$, where ω_1 and ω_2 correspond to the concept for higher-WTP and lower-WTP consumers, respectively. The formulation of the optimization problem now becomes.

Respondent #	248	1	5	15	1	
	5					Number of choice tasks
# of concepts in choice task	3	1				Number of attributes
	2	1	3	2	2	Level combination of concept #1
	3	2	2	3	1	Level combination of concept #2
	5	1	5	4	5	Level combination of concept #3
Concept chosen	1	12				

Figure 3 Sample data of a choice task from a particular respondent

$$\begin{aligned} \max_{(p_1, p_2)} m(p_1, p_2) &= d_1(p_1)(p_1 - c_1) + \\ d_2(p_2)(p_2 - c_2) \quad s. t. \quad &p_1, p_2 \geq 0 \end{aligned} \quad (6)$$

In addition to the nonnegativity constraint, we need to add those representing the interdependence between demands. Suppose we want to determine optimal prices in a two-segment problem and start with initial prices $p_1^{(0)}, p_2^{(0)} \in \{p_1, \dots, p_m\}$. In this situation, $d_1(p_1)$ is derived using RFC simulation under a scenario in which the price for a lower-WTP segment is $p_2^{(0)}$. Referring to the RFC_i function defined in the previous subsection, this can be expressed as

$$d_1(p_1) = f(RFC_1(\mathbf{X}, \boldsymbol{\beta} \mid \{\Omega, \omega_1, \omega_2(p_2^{(0)})\}), D) \quad (7)$$

where $\omega_2(p_2^{(0)})$ refers to a lower-price concept that is offered at the price of $p_2^{(0)}$. For simplicity, we will solve the problem for one segment at a time. We start by solving for the higher-WTP segment. Suppose we find an optimal price of \hat{p}_1^* , then the current solution becomes $(\hat{p}_1^*, p_2^{(0)})$. Now, we want to solve for the lower-WTP segment with the corresponding demand function that can be expressed as

$$d_2(p_2) = f(RFC_2(\mathbf{X}, \boldsymbol{\beta} \mid \{\Omega, \omega_1(\hat{p}_1^*), \omega_2\}), D) \quad (8)$$

where $\omega_1(\hat{p}_1^*)$ refers to a higher-price concept that is offered at the price of \hat{p}_1^* . Suppose we get an optimal price of \hat{p}_2^* , then the current solution is $(\hat{p}_1^*, \hat{p}_2^*)$.

Here comes the complication. Our optimal price for a higher-WTP segment (\hat{p}_1^*) is obtained from an optimization where the corresponding demand function, $d_1(p_1)$, comes from a scenario which assumes that the price for a lower-WTP segment is $p_2^{(0)}$. Now, the price for the lower-WTP segment is \hat{p}_2^* , which means that \hat{p}_1^* may be no longer optimal. Hence, we need to repeat the cycle and find the optimal price for higher-WTP segment under a scenario where the price for a lower-WTP segment is \hat{p}_2^* . Subsequently, we may need to revise the optimal price for the lower-WTP segment, and so on. Another complication is that it is quite possible that \hat{p}_1^* and \hat{p}_2^* in any cycle do not correspond to any of the element in $\{p_1, \dots, p_m\}$, which means that we do not have the information about their utility values. In this paper,

we resolve this by rounding the value of \hat{p}_1^* and \hat{p}_2^* in any cycle to the closest value in set $\{p_1, \dots, p_m\}$, and repeat the cycle until the solution does not change.

This approach can be generalized to pricing with n fare-classes in which $p_1 \geq p_2 \geq \dots \geq p_n$, where p_1, \dots, p_n are prices for fare-class 1, ..., n . The number of fare-class should not exceed the number of levels set for the price attributes of the choice data. Given the set of price levels of the choice data, $\mathcal{P} = \{p_1, \dots, p_m\}$, the approach works as follows:

Step 0

Set initial solution (price levels) of each fare-class $[p_1^{(0)}, \dots, p_n^{(0)}]$ where $p_1^{(0)}, \dots, p_n^{(0)} \in \mathcal{P}$.

Determine demand function of each fare-class, $[d_1(p_1), \dots, d_n(p_n)]$, based on the initial price levels. Use $[d_1(p_1), \dots, d_n(p_n)]$ as the current demand functions and $\mathbf{p} = [p_1^{(0)}, \dots, p_n^{(0)}]$ as the current solution.

Step 1

Set $i = 1$

1. Based on \mathbf{p} , solve for fare-class i to get the current optimal price for fare-class i , \hat{p}_i^* . If $\hat{p}_i^* \in \mathcal{P}$, set \hat{p}_i^* as the current solution for fare-class i and revise \mathbf{p} accordingly. Otherwise, set \hat{p}_i^* to \hat{p}_i^* , the element of \mathcal{P} that is closest to \hat{p}_i^* , and revise \mathbf{p} accordingly to get \mathbf{p}' .
2. If $i = n$, go to Step 2. Otherwise, set $i = i + 1$. Revise the demand function for fare-class i , $d_i(p_i)$, based on \mathbf{p}' . Go to 1.

Step 2

If the solution does not change, stop. Otherwise, go to Step 1.

Results and Discussions

Our case study is pricing optimization for mobile broadband services in the price-sensitive market in Indonesia. The dataset that we use in this research is from Susanta [24], which contain data from 231 respondents. The data was collected through an online survey using a choice-based conjoint (CBC) questionnaire that is designed using attributes and levels as in Table 1.

Table 1. Attributes and levels of mobile broadband services

Attributes	Levels	
	Code	Name
Carrier	1	Telkomsel
	2	Indosat
	3	Axis
	4	Smartfren
	5	Flexi
Payment	1	Prepaid
	2	Postpaid
FUP quota	1	500 MB
	2	2 GB
	3	5 GB
	4	8 GB
	5	10 GB
Maximum speed	1	768 kbps
	2	1.8 Mbps
	3	3.1 Mbps
	4	7.2 Mbps
	5	14.4 Mbps
Price	1	Rp50,000
	2	Rp90,000
	3	Rp130,000
	4	Rp250,000
	5	Rp450,000

Each respondent is given 13 random choice tasks, each of which consists of three concepts and one “none” option. A concept represents a mobile broadband monthly plan with a fixed price. The “none” option makes the questionnaire more realistic since in real-world people may decide not to choose among the available options. Questionnaire for each respondent is unique and randomly generated such that it is balanced and orthogonal. Two fixed choice tasks (same for all respondents) are added for validation purpose.

The raw data is in a format, as shown in Figure 3. For each choice task, respondents choose the plan that they prefer the most. Otherwise, they should choose the “none” option. Each respondent must respond to 15 choice tasks presented in a similar format repeatedly. This may cause boredom which may affect the quality of data. The utility values were calculated from the choice data using HB estimation method using Sawtooth Software CBC/HB version 5.5.6. Table 2 shows the result of HB estimation of sample respondent.

Utility values as in Table 2 are zero-centered, and a higher value represents a preferred level. The fitness value of this respondent is 0.6517, which means that predicting choice for this respondent using utility values in Table 2 is $0.652/0.250 = 2.608$ times better than a random guess. The denominator of 0.250 comes from the fact that there are four alternatives in each choice task, three concepts and one “none” option. The overall fitness of the estimation for all respondents is 0.659.

Table 2. Utility values of sample respondent

Respondent #	Attributes	Levels	Utility values
248	Carrier	Telkomsel	0.9053
		Indosat	0.9125
		Axis	1.5306
		Smartfren	-1.6290
		Flexi	-1.7194
	Payment	Prepaid	0.6917
		Postpaid	-0.6917
	FUP quota	500 MB	-3.1798
		2 GB	-0.8702
		5 GB	0.6397
		8 GB	1.0347
		10 GB	2.3755
	Maximum speed	768 Kbps	-0.4376
		1.8 Mbps	-0.6966
		3.1 Mbps	-0.0475
		7.2 Mbps	0.2287
		14.4 Mbps	0.9530
	Price	Rp50,000	2.8534
		Rp90,000	3.1201
Rp130,000		0.3401	
Rp250,000		-2.2815	
Rp450,000		-4.0321	

This is a moderate value but relatively good considering the sample size of only 231. A typical CBC study usually suggests a minimum sample of 300. We also check the internal validity of the utility model by comparing actual choices with those predicted by the utility values and found a mean absolute error of 6.53%.

For our pricing optimization problem, we pick the Flash Optima plan, which is a prepaid service from Telkomsel with 2GB quota and 7.2Mbps maximum speed, at a monthly price of Rp90.000. We consider a scenario where this plan competes with other prepaid plans as in Table 3. Simulation with randomized first choice using Sawtooth Software SMRT 4.23 produced result in Table 4. We add two price levels in addition to the existing five. These two additional levels are one representing demand when the price falls to 0, and one representing the satiating price (price at which demand falls down to 0).

The maximum achievable market size is assumed to be 50 million. The result indicates a normal product where lower price generates higher demand which corresponds to a downward-sloping demand function.

Since RFC assumes a probabilistic decision making, the share-of-preference estimates come with the corresponding standard error values. The standard errors are relatively small and still consistent with the downward-sloping demand function.

Applying monotonic cubic spline interpolation on the demand data points from the second and fifth column of Table 4, we obtain the spline functions as shown in Table 5.

Table 3. Competitors of our product of interest

Plans	Carriers	FUP quota	Maximum speed	Price
Super Internet	Indosat	500 MB	7.2 Mbps	Rp50.000
Axis Pro	Axis	2 GB	7.2 Mbps	Rp50.000
Connex Evo	Smartfren	8 GB	768 Kbps	Rp90.000
FMB Set	Flexi	5 GB	3.1 Mbps	Rp90.000

Table 4. Randomized first choice simulation result for our product of interest

Plan	Prices	Share-of-preference estimates	Standard error of estimates	Demand
Flash Optima	Rp0	100%		50,000,000
	Rp 50,000	44.71%	2.29%	22,355,000
	Rp 90,000	27.66%	2.00%	13,830,000
	Rp 130,000	18.59%	1.82%	9,295,000
	Rp 250,000	8.28%	1.37%	4,140,000
	Rp 450,000	3.88%	0.98%	1,940,000
	Rp 600,000	0%		0

Tabel 5. The demand functions

Price ranges	Demand functions
$0 \leq p < 50,000$	$d(p) = 56,667,470 + 732.25 p + 4.22 \times 10^{-8} p^3$
$50,000 \leq p < 90,000$	$d(p) = 68,169,990 - 1,442.40 p + 1,38 \times 10^{-2} p^2 - 4.98 \times 10^{-8} p^3$
$90,000 \leq p < 130,000$	$d(p) = 29,941,522 - 148.12 p - 3,56 \times 10^{-4} p^2 + 2.67 \times 10^{-9} p^3$
$130,000 \leq p < 250,000$	$d(p) = 39,679,353 - 372.84 p + 1,37 \times 10^{-3} p^2 - 1.77 \times 10^{-9} p^3$
$250,000 \leq p < 450,000$	$d(p) = 13,968,855 - 64.31 p + 1,39 \times 10^{-4} p^2 - 1.20 \times 10^{-10} p^3$
$450,000 \leq p < 600,000$	$d(p) = -1,698,896 + 40.14 p - 9,33 \times 10^{-5} p^2 + 5.18 \times 10^{-11} p^3$

Tabel 6. Iterations of pricing optimization for scenario 2

Iteration	\hat{p}_1^*	$\hat{p}_1^{*'} $	\hat{p}_2^*	$\hat{p}_2^{*'} $	$m(\hat{p}_1^*, \hat{p}_2^*)$
1	Rp 136,000*	Rp 136,000	Rp 90,000	Rp 90,000	Rp 784,771,751,150
2	Rp 136,000	Rp 130,000	Rp 89,000*	Rp 89,000	Rp 1,027,559,351,020
3	Rp 136,000*	Rp 136,000	Rp 90,000	Rp 90,000	Rp 779,602,526,435
4	Rp 136,000	Rp 130,000	Rp 89,000*	Rp 89,000	Rp 1,027,559,351,020

* indicates optimal price obtained from pricing optimization

The use of cubic splines for demand function brings another complication. If the demand functions in Table 5 goes into the optimization problem as in equation 4, we will have an optimization problem with quartic (4th degree) polynomial objective function. Solving this kind of problem requires an assessment of its convexity. According to Ahmadi *et al.* [1] determining the convexity of a quartic polynomial optimization problem is NP-hard. Hence, using an analytical solution would not be feasible.

In reality, the price is always in multiple of Rp1,000, i.e. {1,000, 2,000, ..., 600,000}. Considering the size of the solution space, we use enumeration to find the solution to our pricing optimization problem. In doing so, we make four optimization scenarios:

1. Pricing optimization for one fare-class
2. Pricing optimization for two fare-classes without product differentiation

3. Pricing optimization for two fare-classes with product differentiation
4. Pricing optimization for three fare-classes with product differentiation

The incremental cost is assumed to be Rp25,000 for one fare-class and two fare-classes without product differentiation, Rp35,000 and Rp25,000 for two fare-class with product differentiation, and Rp45,000, Rp35000, and Rp25,000 for three-class optimization. Product differentiation is limited to be based on one attribute, i.e. FUP quota. The first scenario is used as a comparison to determine if the price differentiation is justified. The demand function for scenario 1 is in Table5, while for other scenarios are derived using the same approach. There are two, two, and three demand functions in scenario 2, scenario 3, and scenario 4, respectively.

Tabel 7. Iterations of pricing optimization for scenario 3

Iterasi	\hat{p}_1^*	\hat{p}_1^{*} '	\hat{p}_2^*	\hat{p}_2^{*} '	$m(\hat{p}_1^*, \hat{p}_2^*)$
1	Rp 124,000*	Rp 124,000	Rp 90,000	Rp 90,000	Rp 1,479,323,846,136
2	Rp 124,000	Rp 130,000	Rp 89,000*	Rp 89,000	Rp 1,518,372,036,157
3	Rp 124,000*	Rp 124,000	Rp 89,000	Rp 90,000	Rp 1,474,916,919,729
4	Rp 124,000	Rp 130,000	Rp 89,000*	Rp 89,000	Rp 1,518,372,036,157

* indicates optimal price obtained from pricing optimization

Tabel 8. Iterations of pricing optimization for scenario 3

	\hat{p}_1^*	\hat{p}_1^{*} '	\hat{p}_2^*	\hat{p}_2^{*} '	\hat{p}_3^*	\hat{p}_3^{*} '	$m(\hat{p}_1^*, \hat{p}_2^*, \hat{p}_3^*)$
1	Rp 146,000*	Rp 146,000	Rp 90,000	Rp 90,000	Rp 50,000	Rp 50,000	Rp 1,399,870,394,858
2	Rp 146,000	Rp 130,000	Rp 99,000*	Rp 99,000	Rp 50,000	Rp 50,000	Rp 1,550,077,237,849
3	Rp 146,000	Rp 130,000	Rp 99,000	Rp 90,000	Rp 89,000*	Rp 89,000	Rp 1,985,004,854,271
4	Rp 145,000*	Rp 145,000	Rp 99,000	Rp 90,000	Rp 89,000	Rp 90,000	Rp 1,565,141,602,715
5	Rp 145,000	Rp 130,000	Rp 99,000*	Rp 99,000	Rp 89,000	Rp 90,000	Rp 1,698,771,803,345
6	Rp 145,000	Rp 130,000	Rp 99,000	Rp 90,000	Rp 89,000*	Rp 89,000	Rp 1,974,144,366,274
7	Rp 145,000*	Rp 145,000	Rp 99,000	Rp 90,000	Rp 89,000	Rp 90,000	Rp 1,565,141,602,715

* indicates optimal price obtained from pricing optimization

Scenario 1 results in the optimal price for Flash Optima of Rp118,000 with share-of-preference of 20.48% with a total contribution of Rp 952,410,781,139. In scenario 2, Flash Optima is offered at two different prices but with exactly the same non-price attribute levels. The initial prices are Rp 130,000 and Rp 90,000. The optimization result from all iterations is shown in Table 6.

In iteration 1, we solve for the fare-class 1, assuming that the price for fare-class 2 is Rp90,000. We found that the optimal price for fare-class 1 is Rp136,000. In iteration 2, we solve for fare-class 2. Since the optimal price for fare-class 1 (Rp 136,000) is not the element of the price levels defined for the choice data, the demand function in this iteration is derived by assuming that the price for fare-class 1 is Rp 130,000 (level closest to Rp 136,000). This results in optimal price of Rp89,000 with significant increase in total contribution from Rp 784,771,751,150 to Rp 1,027,559,351,020. This significant difference is due to different price levels used for estimating demand and calculating total contributions for fare-class 1. Demand is estimated based on \hat{p}_1^{*} ' = Rp130,000 (which results in greater demand than using the optimal price of Rp 136,000), while total contribution is calculated using \hat{p}_1^* = Rp 136,000. Using similar reasoning, we can infer that the total contribution of Rp779,602,526,435 in iteration 3 is less than it should, and the total contribution of Rp 1,027,559,351,020 in iteration 4 is greater than it should. We can also observe that the solution stays at [Rp 136,000, Rp 89,000]. Then, we can infer that under optimal pricing of [Rp 136,000, Rp89,000] the total contribution is between Rp 779,602,526,435 and Rp 1,027,559,351,020. From the pricing optimization with one fare-class, we have a total contribution of Rp 952,410,781,139. Hence, we cannot conclude that

price differentiation without product differentiation with two fare-classes is better than single-fare policy, or vice versa. Solving scenario 2 under independent demand assumption using equation 5 results in optimal prices of \hat{p}_1^* =Rp406,000 and \hat{p}_2^* =Rp 103,000 with total contribution of Rp 1,629,189,562,533. Hence, compared to the result under interdependent demand, cannibalization dilutes total contribution by at least 37%.

In the third scenario, Flash Optima is offered at two different options, i.e. with 5GB quota offered initially at Rp 130,000 and 2GB quota initially offered at Rp 90,000. The result of the pricing optimization is shown in Table 7.

This optimization results in optimal prices of [Rp 124,000, Rp 89,000] which produces total contribution of Rp 1,518,372,036,157. Considering that this figure is based on the demand function which assumes p_1^{*} ' = Rp 130,000 and p_2^{*} ' = Rp90,000, we can infer that the true value of total contribution is greater than Rp 1,518,372,036,157. Compared to the result from the first scenario, price differentiation with product differentiation results in an almost 60% increase in total contribution.

In general, more fare-classes will result in greater total contribution in the absence of cannibalization. In the last scenario, we investigate if adding one more fare-class will increase the total contribution. We set three fare-classes, i.e. 8GB initially offered at Rp130,000, 5GB initially offered at Rp 90,000, and 2GB initially offered at Rp 50,000. The iterations of the pricing optimization for scenario 3 are shown in Table 8. After seven iterations, it was found that the optimal pricing policy for this scenario is [Rp145,000,

Rp 99,000, Rp 89,999]. From iteration 3 it can be inferred that the price for fare-class 3 is too low. Starting in iteration 3, the price level for fare-class 2 and 3 in the RFC simulation, based on which the demand functions are derived, are the same, i.e. Rp90,000. From iteration 3 to 7 it can be inferred that the total contribution under optimal pricing is less than Rp1.565.141.602.715. From scenario 3 we know that total contribution for two fare-classes with product differentiation is greater than Rp 1,518,372, 036,157. Hence, we cannot conclude that three fare-classes is better than two. In reality, managing three fare-classes is more difficult than two. Hence, in this case study, it is recommended to implement price differentiation with product differentiation with two fare-classes.

Conclusions

We develop an approach to incorporate cannibalization into pricing optimization using choice data. From choice data, individual utilities are estimated using hierarchical Bayes. Based on these utility values, individual choice in a particular choice situation is estimated using randomized first choice simulation. Using this mechanism, demand function under any competitive scenario can be estimated. Cannibalization is modeled by incorporating the cannibalizing products into the scenario based on which the demand function is derived. Determining optimal prices for cannibalizing products involves optimization to maximize the total contribution of multiple fare-classes with intertwined demand functions. The fixed price levels defined for the choice data adds more complication to the problem. We propose an iterative procedure to solve the problem and implement it in the pricing of mobile broadband services in Indonesia. For our case study, it is shown that a two-fare-class price differentiation with product differentiation is better than a single-price policy and three fare-class policy. The two-fare-class policy results in about 60% increase in total contribution compared to the single-price policy.

While proved to be effective, some issues regarding our proposed approach still need to be addressed, e.g. resolving the problem caused by the fixed price levels, and coping with problems with many fare-classes like those in the airline industry. In the future, research should be directed toward those areas.

References

1. Phillips, R. L., *Pricing and Revenue Optimization*, Stanford University Press, Stanford, 2005.
2. Lomax, W., Hammond, K., East, R., and Clemente, M., The Measurement of Cannibalization, *Journal of Product & Brand Management* 6(1), 1997, pp.27-39.
3. Carpenter, G. S., and Hanssens, D., M., Market Expansion, Cannibalisation, and International Airline Pricing Strategy, *International Journal of Forecasting* 10, 1994, pp.313-326.
4. Meredith, L., and Maki, D., Product Cannibalization and the Role of Prices, *Applied Economics* 33, 2001, pp.1785-1793.
5. Srinivasan, S. R., Ramakrishnan, S., and Grasman, S. E., Incorporating Cannibalization Models into Demand Forecasting, *Marketing Intelligence & Planning* 23(5), 2005, pp.470-485.
6. Albuquerque, P., and Bronnenberg, B. J., Estimating Demand Heterogeneity Using Aggregated Data: An Application to the Frozen Pizza Category, *Marketing Science* 28(2), 2009, pp.356-372.
7. Heerde, H. J. v., Srinivasan, S., Dekimpe, M. G., Estimating Cannibalization Rates for Pioneering Innovations, *Marketing Science* 29(6), November-December 2010, pp. 1024-1039.
8. Pacras, J., Sriram, S., and Kumar, V., Empirical Investigation of Retail Expansion and Cannibalization in a Dynamic Environment, *Management Science* 58(11), 2012, pp.2001-2018.
9. Haynes, M., Thompson, S., and Wright, P. W., New Model Introductions, Cannibalization and Market Stealing: Evidence from shopbot data, *Manchester School* 82(4), 2014, pp.385-408.
10. Ngwe, D., Why Outlet Stores Exist: Averting Cannibalization in Product Line Extensions, *Marketing Science* 36(4), 2017, pp.524-541.
11. Guo, Z. and Chen, J., Multi-Generation Product Diffusion in the Presence of Strategic Customers, *Information Systems Research* 29(1), 2018, pp. 206-224.
12. Reimers, I., and Xie, C., Do Coupons Expand or Cannibalize Revenue? Evidence from an e-Market. *Management Science* 65(1), January 2019, pp. 286-300.
13. Carpenter, G. S., and Hanssens, D., M., Market Expansion, Cannibalisation, and International Airline Pricing Strategy, *International Journal of Forecasting* 10, 1994, pp.313-326.
14. McFadden, D., Modelling the Choice of Residential Location. In: Snickars, F., Weibull, J. (Eds.), *Spatial Interaction Theory and Planning Models*. North-Holland, 1978, pp.72-77.
15. Allenby, G. M., Arora, N., and Ginter, J. L., Incorporating Prior Knowledge into the Analysis of Conjoint Studies, *Journal of Marketing Research* 32(2), 1995, pp.152-162.
16. Allenby, G. M., and Ginter, J. L., Using Extremes to Design Products and Segment Markets, *Journal of Marketing Research* 32(4), 1995, pp.392-403.
17. Huber, J., Orme, B., & Miller, R., Dealing with Product Similarity in Conjoint Simulations. In A. Gustafsson, A. Herrmann & F. Huber (Eds.), *Conjoint Measurement: Methods and Applications* (4th ed.), Springer, 2007, pp. 347-362.

18. Train, K. E., *Discrete Choice Methods with Simulation*, 2nd edition, Cambridge University Press, 2009.
19. Sawtooth Software, The CBC/HB System for Hierarchical Bayes Estimation Version 5.0 Technical Paper, *Technical Paper Series*, 2009.
20. Huber, J., Orme, B., & Miller, R., Dealing with Product Similarity in Conjoint Simulations. In A. Gustafsson, A. Herrmann & F. Huber (Eds.), *Conjoint Measurement: Methods and Applications* (4th ed.), Springer, 2007, pp. 347–362.
21. Huang, J., Leng, M., and Parlat, M., Demand Functions in Decision Modeling: A Comprehensive Survey and Research Directions, *Decision Sciences*, 44(3), 2013, pp. 557-609.
22. Wolberg, G. and Alfy, I., An Energy-Minimization Framework for Monotonic Cubic Spline Interpolation, *Journal of Computational and Applied Mathematics*, 143, 2002, pp. 145-188.
23. Pratikto, F. R., Uncapacitated Pricing Optimization for Mobile Broadband Services, *Jurnal Teknik Industri* 20(1), June 2018, pp. 49-58.
24. Susanta, L. N., *Usulan Penetapan Harga untuk Paket Mobile Internet Telkomsel Flash Berdasarkan Choise-Based Conjoint Analysis*, Final Project, Department of Industrial Engineering, Parahyangan Catholic University, 2013.