Estimating Price Elasticity of Demand for Consumer Credit in Automated Online Environment Using SAS

Liyan Zhao, Feng Xu

Abstract:

Two measures of price elasticity of demand for consumer credit, purchase elasticity of price and balance elasticity of price, are graphically analyzed and statistically estimated using random price test data. Key business impacts are simulated. Estimation and simulation are coded in SAS software and can be used in periodically updating the estimation and simulation when longer performance of price test data become available in helping determine the optimal pricing for credit lending.

Introduction

Oftentimes price elasticity (PE) is practically not used in price determination for retail credit products. It is actually pretty common that an organization simply ignores the PE by just setting up one price to fit all, or setting up prices to target certain threshold margin. There could be multiple reasons for this neglect. On one extreme, some decision makers simply do not believe the existence of PE or price is very inelastic. More likely reasons are lack of empirical PE results. Price tests are usually very expensive and many times the infrastructure even prevents from setting up valid price tests in many organizations.

This paper provides empirical results of PE for automated online sales. The major questions addressed are: whether the PE exists, what the magnitude of PE is, and how the trade-off between price and credit demand can be used to determine the best or optimal pricing. This paper first outlines the price test set-up, then discusses the data organization, graphical analyses of PE, statistical estimation of PE, and impact simulations of price, all using SAS software, and finally summarizes the major results.

Price Test and Data

Among Dell buyers who have applied for credit to finance their Dell purchase from online channel, portion of them are approved for credit by an automated decision system. An approved applicant is offered a price that this applicant can accept or decline.

Approved applicants in this analysis are limited to online channel applicants who are not qualified for individual level promotion. The approved applicants are first divided into several unique sub-populations based on their risk profile, let the number of sub-populations be **m**. Each sub-population is then randomly divided into test cells, let the number of cells in each subpopulation be **n**. Total number of population*cell combinations is m*n, each combination has offered a unique price.

Price test was conducted in a 80-day period from June 21 to September 10, 2007. Performance period up to June 2008 was first analyzed using SAS and the SAS codes were subsequently used to analyze the performance up to July 2008 when the new data became available. The following data were gathered: Whether an approved applicant accepted the price offer and made a purchase; If accepted, the life-to-date (ltd) purchase amount, end balance as of July 2008, ltd PnL items including interest, fees, cost of funds, charge-off, operation cost.

Data Analysis and PE Estimation

Price impact measurements and PE's are defined as below. <u>Take elasticity of price, TEP</u>, measures take (accept offer and make purchase) behavior in response to price. It is hypothesized that more customers will accept a lower price to make purchase than a higher price. Purchase elasticity of price, PEP, measures ltd purchase behavior in response to price. It is hypothesized that customers will purchase more at lower price, due to more approved customers makes purchase (TPE) and/or per customer buys more at lower price. Balance elasticity of price, BEP, measure 'end of measurement period' balance behavior in response to price. It is hypothesized that low price will have more remaining balances than high price due to competitiveness in price. Balance is a major source of revenue.

Graphic analysis and results. Take rate, per approve ltd purchase, and end balance are graphed against price, for each sub-population. Graphic analyses show that take rate, purchase per approve, and balance per approve decrease as price increases within each sub-population, as one would expect directionally. The graphic results intuitively supports the hypotheses listed above.

Equation estimation and results. Only purchase and balance elasticities are statistically estimated since take impact is implicitly included in these two measures.

SAS is used to estimate PEP and BEP. From graphics results, the slopes of purchase*price is very similar among different sub-populations, the same is true for balance*price relationship among sub-populations. Instead of estimate an equation for each sub-populations, all subpopulations are pooled to estimate one equation with the differences in intercept are dummied. Independent variables also included days of performance, weekday dummies. Model specifications are as below:

LTD purchase for each approved applicant = a function of:
Offer price of each individual applicant,

Sub-population dummies for various risk profiles,
Days from decision date to last performance date,
Weekday dummies of decision dates.

An equation can also be estimated for each sup-population separately. Since price is the only the variant within each sub-population, ‘noise’ in data for estimating PEP and BEP is better controlled. However, there will be multi-equations. Instead, PE is estimated for all the sub-populations together, with the impact of each sub-population is dummied. This method has overcome the difficulties of small sample issue for some test cells and sub-populations which is the case in our price test data. Also, the graphic results indicate the similar slope among sub-populations. Each method has its advantages. When test data permitting, method one is preferred as a staring point. Vigorous statistical tests may support multi-equation over single equation or vice versa. Oftentimes, single equation has to be used due to test data in a business environment. Estimation results indicate that two PEs are statistically significant.

Impact Simulation

Four scenarios are analyzed: highest price for each sub-population, let it be the baseline for the analysis purpose here, and 150, 350, and 550 bps price reductions cross the board. Key business measures such as the number of purchase customers, total sales, total balances, major PnL (profit and loss) items are all simulated for various price*sub-population combinations using SAS macros.

The optimal pricing can be set up using a mathematical integer programming framework. Instead of setting up the mathematical programming framework, exhaustive combinations of price*sub-population are set up in SAS as the dimensionality is manageable using SAS macros. For example, for seven sub-populations, each has four possible prices, there are a total number of 16,384 sub-population*price combinations (4^^7).

Simulation results help answer various business questions. For example, relative to the key business measures from the Baseline price, what are the impact of various of price changes. In order to increase the sales by 10%, how should the prices be changed: price reduction by xx bps in sub-population 1, price reduction by yy bps in sub-population 2, no price change in sub-population 3, etc. Also there are price change opportunities for higher profitability at the same sales level.

What will happen if price is reduced to 550 bps cross the board, sales gain, balance gain, interest revenue reduction, and profit loss? Simulation results can help depict various frontiers (the best combinations of y*x) such as the sales-price frontier, balance-price frontier, and profit-price frontier. Inefficient pricing is found to be quite significant in some measures. Inefficiency is defined as the purchase (or balance) difference away from the optimal level at the same price.

Summary

Online channel price elasticities are graphically analyzed and statistically estimated for credit lending. Key business measure impacts of various price*population combinations are simulated. SAS software is used in estimation and simulation. The analytical framework using coded SAS macros will be used in periodically updating the estimation and simulation when longer performance of price test data becomes available in helping determine the optimal pricing for credit lending.

SCSUG2008
END

Appendix (1)

Purchase elasticity of price: measures the changes in purchase amount in response to price. Per approve purchase amount is affected by take rate and per take purchase amount (both initial purchase and repeat purchase) from different price points.

Hypothesis:

- Purchase elasticity of price exists: a1 ne 0
- Purchase elasticity of price exists: a1 > 0
- Purchase elasticity of price exists: 95%CI.low < a1 < 95%CI.high

Model specification for Online channel only and non-alp only:
Log (per approve ltd as of June 2008 purchase amt $) =
+ a0
+ a1 * log(price) [a1 is the elasticity]
+ b3 * pb3_ind
+ b4 * pb4_ind
+ b5 * pb5_ind
+ b6 * pb6_ind
+ b7 * pb7_ind
+ b8 * pb8_ind
+ c1 * ARS (line or log form)
+ d1 * Days between application date and the last day of performance period June 2008 (line or log form)
+ other variables

Appendix (2)

Balance elasticity of price: measures the changes in remaining balance in response to price. Per approve balance amount is affected by purchase and payment behaviors from different price points.

Hypothesis:

- Balance elasticity of price exists: a1 ne 0
- Balance elasticity of price exists: a1 > 0
- Balance elasticity of price exists: 95%CI.low < a1 < 95%CI.high

Log (per approve mth 9 balance $) =
a0
+ a1 * log(price) [a1 is the elasticity]
+ b3 * pb3_ind
+ b4 * pb4_ind
+ b5 * pb5_ind
+ b6 * pb6_ind
+ b7 * pb7_ind
+ b8 * pb8_ind
+ c1 * ARS (line or log form)
+ d1 * ltd_purchase (line or log form)
+ e1* Days between application date and the last day of performance period June 2008 (line or log form)
+ other variables

Optimal Price at Individual Level:

Population selection criteria

1. Excluding Offline Channel
2. Choose a particular risk segment. The highest risk segment is chosen in the case study here.

Procedures

1. Develop four take score models (logistic models)
    a. T2474 = f(credit bureau, ars, application data)
    b. T2699 = f(credit bureau, ars, application data)
    c. T2824 = f(credit bureau, ars, application data)
    d. T2999 = f(credit bureau, ars, application data)
2. Score ~35,000 approved applicants using the four models. Expected results are
    a. Mean take score T2474 > T2674 (2474+200) > T2824 (2474+350) > T2999 (2474+525). In other words, models should estimate a higher take rate for 2474 than for 2999, for the ~35,000 approved applicants.
    b. It is hypothesized that some individuals are more sensitive to price than others. Use T2999 and T2824 as an example. Difference between 2999 and 2824 = T 2999 – T2824. Sort the difference from most positive to most negative. The following groups can be generated:
        i. Very positive, i.e. T2999 is much higher than T2824. Only a few customers are expected to be in this category. They are 'irrational' customers who are estimated to be more responsive to 2999 than to 2874. What are the actual test results? If

actual test results support the estimation, they are good candidates for price at 2999.

ii. Somewhat positive, i.e. T2999 is slightly higher than T2824. Again only a few customers are expected to be in this category. They are ‘irrational’ customers who are estimated to be slightly more responsive to 2999 than to 2874. What are the actual test results? If actual test results support the estimation, they are good candidates for price at 2999.
iii. Somewhat negative, i.e. T2999 is slightly lower than T2824. They are estimated to be slightly less responsive to 2999 than to 2874. What are the actual test results? If actual test results support the estimation, they are good candidates for price at 2999.
iv. Very negative, i.e. T2999 is much lower than T2824. They are estimated to be much less responsive to 2999 than to 2874. What are the actual test results? If actual test results support the estimation, they are good candidates for price at 2874.
v. i-iv are for illustrative purpose only. Simulation should help decide the number of groups and break scores for the two-score case. When number of scores increases, more complex simulation is needed.