Pricing Analytics Project 2:

Kiwi Bubbles

Takeaki Sunada

1 General notes

Please send your answers by Friday, Feb 14 by the end of day (11:59pm), to "mkt440submission@gmail.com". The submission should take either of the following forms.

- Option 1 (3 files to attach): A write-up in either pdf or html format (one per team), your .R code and an HTML file compiled by Rstudio. The HTML file is generated by selecting "files - compile report" in Rstudio console. It then automatically runs your R code and report all results side by side with your code. In this case, your write-up should be in a separate file and not mixed up in the code.

- Option 2 (1 file to attach): An R-markdown report, either in pdf or in html format, that puts together your code, results and the write-up. Please put your write-up as a text in the report, and not as a comment in the code.

Either case, please use as the file name "first names of the team members, in alphabetical order, separated by underscore (_)". No need to say "Project 2".

In your write-up, please clearly indicate the first and last names and student ID of all team members. Also please clearly label each of your answers (for example, "answer to question 4-1"), so that we know which question you are addressing. I will let our graders subtract points if formatting/labeling issues cause them hardship to grade. Concise write-ups are greatly appreciated (though of course we understand sometimes it just takes space to justify your specifications).

2 Environment

It is late 90s. You are the product and pricing manager of Kiwi, a major soft drink company. You own and sell "Kiwi Regular" and the product currently is doing well in the market. Your main competitor, Mango, produces "Mango Bubbles" that is well-received in the market as well. Your guess is that Mango does well because consumers prefer "Bubbles" over "Kiwi" – but you have the better brand name than Mango, thus your product also does well. Now, a spark comes to your mind: why not launching a new product, "Kiwi Bubbles"? The product combines the best of characteristics "Bubbles" and your brand name "Kiwi", and should be able to do well, right? You launch the product in a test market and collected consumer loyalty card data to aid your global launching and pricing decisions. Throughout the project, assume that all 3 products have $0.50 unit costs and market size is 1000 consumers. We have a subsample of the data set, "kiwi_bubbles_P2.csv" (note that this is *a different data set* from the one we use in the lecture). The data set records the choices of soft drink by randomly sampled 359 consumers over the course of 3 years. The columns, from left to right, are 1) consumer ID, 2) week code, 3) softdrink-buying trip number of the consumer in the sample period, and 4-7) prices of zero product (not buying), Kiwi Bubbles, Kiwi Regular and Mango Bubbles (99 means that product was out of stock), and 8) consumer choice. A consumer in a given week only buys one or none of the 3 products. We also have a demographic data-set, "demo.csv", for 283 consumers (whose rough description is available in "demo_description.xlsx").

Before you start - several pieces of advice

1. For many of the questions, you can reuse your codes from the previous questions and save (lots of) time. For example, you need to solve Kiwi and Mango's optimal pricing problem multiple times in Section 5. This is a simple iteration (with minor edits) of Kiwi's optimal pricing you solve in Section 4-4.

2. The latter part of the project asks you to discuss positioning of Kiwi Bubbles. If you don't find anything useful for that question in your estimated demand, you may have the segmentation incorrect - try a different number of segments.

3. Demographic variables are categorical and often the number itself doesn't mean anything

(e.g., occupation = 4 is not any "closer" to occupation = 3 than occupation = 10). Hence in principle, running K-mean clustering on these variables can be somewhat strange. For this project, it is fine to ignore this issue and feed all the variables into Kmean function.

4. When you choose the N in various parts of the project (like when you compute segment share, or when you predict profits), think of what N you should use. The market size is 1000 people, of which 359 people are in the data, of which we see demographics of 283 people. Some parts should use $N = 1000$ and other parts should use $N = 359$.

5. The use of generative AI is allowed. However, for estimation purposes, I *strongly* recommend making use of our template code and lecture codes and manually giving necessary edits. Generative AIs should be limited to producing figures, etc. If you ask AI to estimate the model for you, the AI tends to pull some random packages (that I don't even know) and code up some random (perhaps looking similar to our model, but not quite exact) stuff. Fixing that is much harder than editing the template codes. As usual, if any parts of the project make use of generative AIs, please properly acknowledge it in your report (i.e., explain where and how it is used).

3 Logit model without segmentation

Given that we have a choice data set, it is natural to start with a logit model. Assume first that β_0^j and β_1 are common across consumers. In this case, the model collapses to a simple multinomial logit model. Please feel free to make use of the template code, section 1.

Questions:

1. Estimate the model and report the estimates. What do those coefficients mean?

2. Using the estimated parameters, calculate own- and cross-price elasticities for all combination of products, evaluated at the average prices observed in the data. Do you see any particular patterns? Are the patterns reasonable, or something you are concerned about? Explain why (why not).

    - Tip: As a convention, own elasticity can be defined either as a positive or a negative number (the formula on the slide gives you a positive number). Either answer is fine.

3. Calculate optimal prices for KB and KR (note that you have two products) to maximize the profit when Mango price is $P^{MB} = 1.43$.

4 Logit model with segmentation

Now allow β_0^j and β_1 to differ across consumers with different demographics. For this exercise, focus on the approach with K-mean clustering. The goal of this section is to understand (1) the strategic importance of launching Kiwi Bubbles, and (2) optimal pricing of KB and KR. Please make use of the template code, section 2.

Questions:

1. Like what we did in the lecture, group consumers into segments using "kmeans" function. You may use all columns from the demographic data to do clustering. Then estimate multinomial logit model separately for each segment of consumers. I would recommend number of clusters larger than 7 to find meaningful segmentation (but not too large - if you have too many segments you'd start seeing a segment with like only two people in it. Stop adding segments before you get there). Report your estimates.

2. At the average prices observed in the data, what are the market-level (aggregated across segments) own- and cross- elasticities among these products? How does it differ from the no-segmentation case? Which products are closer substitutes and which products are not as close substitutes?

3. Discuss the preference of the segments you found (who prefers KB over KR, who is more price sensitive, etc.). How does the underlying customer segmentation explain the substitution pattern you see in the elasticity? From the elasticity and underlying segmentation you found, where (i.e. which segment(s)) do you suggest that Kiwi Bubbles should be positioned?

4. To justify your suggestion, compare profits with and without Kiwi Bubbles. Assume for now that Mango Bubbles is priced at \$1.43 and does not react to Kiwi's pricing. First, assume that you decided not to launch Kiwi Bubbles. What is the optimal price of Kiwi Regular, and what is Kiwi and Mango's profit? Next, assume that you do launch Kiwi Bubbles. What are the optimal prices for Kiwi Regular and Kiwi Bubbles? How does the profit of Kiwi and

Mango change as Kiwi launches KB? Does your result justify the launch of KB, and if so, does it support your positioning suggestions in the previous question?

- Tip: when calculating profit without Kiwi bubbles, assume that the preference parameters for KR and MB you estimated stay the same even if you take away KB from the market. In other words, to study profit without KB, you need a multinomial logit model with only KR and MB in the market, and that the preference parameters for those products are identical to what you estimated before. No need to reestimate your no-KB model with the data.

5 Understanding strategic responses

You now consider strategic actions from your competitor, Mango. You realize that Mango might be as smart as you are in that they might also realize their $1.43 price might not be optimal. This is to say, in the previous question, you assumed that Mango maintains its price at $1.43 regardless of what you do. Suppose now instead that they react and set an optimal price against your new prices, and you need to set an optimal prices against their new price, and so on. You wonder where this "pricing war" converge at, and how does your earlier findings about the strategic advantage of launching Kiwi Bubbles changes (or stays the same) under this new scenario. Find them as follows:

Questions:

1. First, solve Mango's optimal pricing problem, given that Kiwi's price is the one you set from the previous section. What is the new price of MB?

2. As Kiwi, you need to react to Mango's new price. Set prices for KR and KB to respond to the new price of Mango Bubble that we just derived. What is your new price for KR and KB?

3. Repeat the previous two steps iteratively, until neither Kiwi nor Mango has an incentive to set a different price (you can be as accurate as one cent, but no need to be more accurate than that). These set of prices are the new "equilibrium price" where you and Mango compete with each other. Please report the sequence of prices until you reach here.

    - Tip: This iterative process should converge within 4-5 iterations (compute optimal prices

for Kiwi and Mango 4-5 times each). Hence you may just copy and paste the previous codes you used to calculate optimal prices four times - conceptually, you just do the same thing repeatedly with different rival price as a starting value. If you don't like such a lengthy code, you may instead write the whole process of deriving optimal prices for each company as a function, and you just call that function at each iteration.

4. At these prices, how does the strategic advantage of Kiwi Bubbles change from the one you derived in the previous section (if at all)?