MKT 440 Pricing Analytics

Takeaki Sunada[1]

[1]Simon Business School
University of Rochester

What is pricing analytics?

- Pricing analytics = "apply data analytic tools to help finding profit-maximizing prices".
- But how does the data help us setting prices? i.e. What do we need to measure from data, in order to set the correct price?

What is pricing analytics?

- Now we have an answer to this question: the data let us see the structure of the demand in the market.
- Most importantly:
  - How consumers' purchase patterns change as we change prices (causal effect of price on demand)
  - How the demand response to prices varies across different segments of consumers (heterogeneity).

Data analytics is all about combining data and models

- Of course, data are incomplete - we need a model to supplement any missing pieces of information.
- However, assumptions imposed on models are ad-hoc and are not often testable (e.g. exponential ratio in logit models). Hence we need a "just right" set of assumptions that allows us to get insights we need, but does not make the model too restrictive.
- What levels of assumptions are "just right" depends on three factors: environment studied, data sets available and managerial objectives.

What environment we study

- Our modeling choices depend heavily on what environment we want to study. By now, we have learned a sufficiently rich set of models to study many environments.
- Demand for baby diaper: demographic variables (kid's age, how many kids the household already had before, etc) should be a good predictor of demand. Moreover, we may want to offer targeted pricing (e.g. offering discounts to first-time parents). Hence demographic-based segmentation may work perfectly.

What environment we study

- Demand for consumer packaged goods (CPG): each firm has a lot of varieties (product name, size, flavor, etc). Hence we may need segmentation analysis for positioning. However demographics likely cannot predict demand well - we may use latent type models. We usually have a long history of choice data to make latent type models work.
- Demand for housing: we need attribute-based models (the number of bedrooms, etc.) to do the cross-product inference. They also provide other useful metrics such as each consumer's WTP for an extra bedroom. Also, demographics should predict well the house each household chooses.

What data we have

- Different data requires different models. Better data requires less assumptions.
- With sales data, a big missing piece in the data is "causal variation". We need many assumptions to complement it (how X variables absorb contamination, conditions for a variable to serve as a valid IV, etc).
- Usually, causal variations are at best a small fraction of all variations in the data. Establishing causality is equivalent to relying on that small set of variations. Unless we have a huge data set, we tend to face some sample size issue (high standard errors of the estimates).
- As a result, we often don't have much data variations to spare to explore other things (e.g., competitive analysis).

What data we have

- With choice data, less concern about contamination - we can use the observed variations for other purposes. Make our models more flexible (i.e. complicated), and we are fine. Moreover, we may see more variables than in a sales-data environment (e.g. demographics).
- Nevertheless, we always need to consider sample size issue. We cannot estimate an overly flexible model relative to our sample size. Otherwise our parameter estimates are not reliable (e.g. Kmeans with too many segments).

What managerial questions we are after

- Different managerial objectives require different models.
- In order to set a single price, estimating models without segmentation would suffice, as we only need average demand in the market (we saw incorporating segmentation won't increase profit much).
- If we want to do targeting or positioning, then we need to identify consumer heterogeneity. We need to assume how heterogeneity unfolds in the market, either through demographic variables, or through latent segments.

Concluding remarks

- Data analytics, including pricing analytics we studied, is all about combining right models with right data. What we all should seek for is the right set of assumptions - can we estimate the effect we want with the data we have?
- If we have too few assumptions, the number we get from our estimates is deemed to be incorrect. However, adding too many assumptions makes the model inflexible, leaving us with an unreasonable estimate.

Concluding remarks

- In this course, we just had a quick glance of the marketing models out there - in the future, you likely encounter various models we have not covered, like "duration model", "buy-till-you-die model", "diffusion model", etc.
- There are simply way too many models to memorize (of course a very popular model, like logit, is probably worth memorizing). Hence, the important skill is to recognize the logic behind each model: "what assumptions does the model use?" "How appropriate those assumptions are in our current environments?". These questions should help you make the right modelling choice in the abundance of models.

Concluding remarks

- Thank you for studying pricing analytics with us! I hope this course has (ever so slightly) changed the way you look at the data.
- Please kindly fill out a teaching evaluation form - this course has been improving thanks to feedbacks from previous years' students. Even if you are completely satisfied with the course (fantastic!), please fill out one - I need the silent majority to speak up! If you are not satisfied (...) PLEASE fill out one - your voice will help next year's students!