ORIE 5355: People, Data, & Systems

Lecture 9: Algorithmic pricing: capacity, price differentiation, and competition

Nikhil Garg

Course webpage: https://orie5355.github.io/Fall_2021/

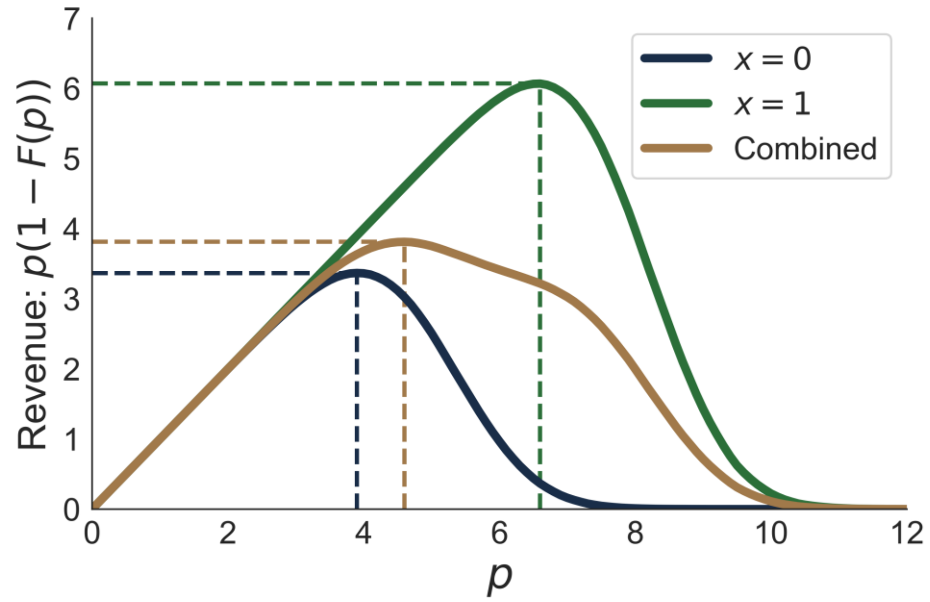
Announcements

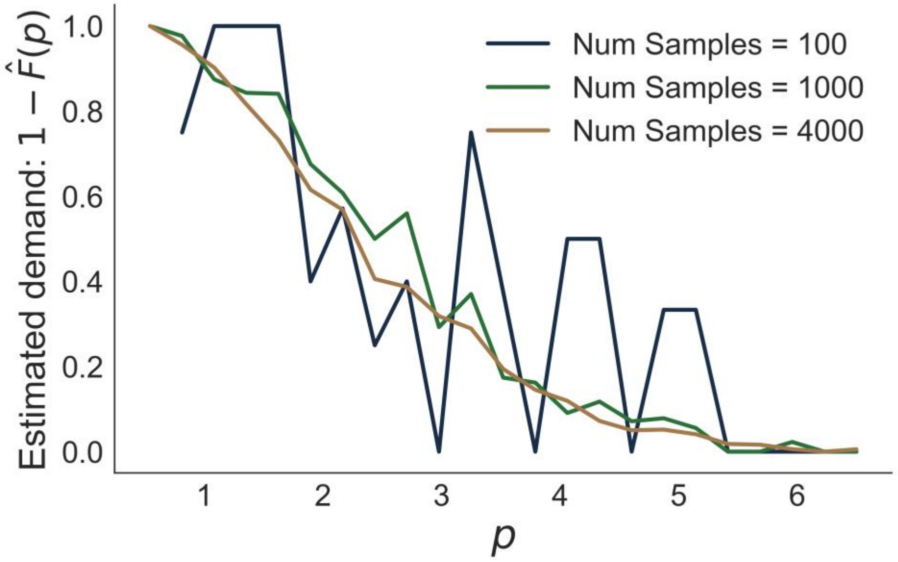
- Quiz 2 released – due Sunday evening
- HW 3 released – due 10/19
- No class on Monday (Fall break)

Last time

- Given a demand distribution $d(p) = 1 - F(p)$, how to calculate optimal prices
$$\arg\max_{p}\,[p \times d(p)]$$

- How to estimate demand distributions, potentially as a function of covariates





More on demand estimation

- We want to estimate $\mathrm{d(p,x)} \stackrel{\text{def}}{=} 1 - F_{p|X}\,(p \mid X = x)$
- Last time: Logistic regression
  - Target (Y variable) is purchase decision $\mathrm{d(p,x)}$
  - Covariates $\mathrm{(p,x)}$ are: price offered, user covariates, interactions between price and covariates or between covariates
- Challenge: what if you have many items you're selling (separately)? This wastes information (can't use models across items)
- Alternative: Use idea from recommendations! Suppose you have user vector $\mathrm{u_i}$ and item vector $\mathrm{w_j}$. Then, ML model to learn with covariates: $(\mathrm{p}, u_i \cdot w_j)$
  - Can learn demand for items you haven't sold before at certain prices!
  - (Or completely new items, using KNN approach from recommendations)
  - Allows incorporating other information you have about items, that helped you learn the item vectors

Plan for today

Many assumptions last time:

- No capacity constraints
- No competition from other sellers
- Only one item
- Allowed to explicitly give different prices to different users
- No over-time dynamics

We’ll peel back some of these assumptions today

Capacity constraints and pricing over time

Setting and examples

You often are trying to sell limited quantities of a good, to many potential customers over time

- Airline tickets – the airline “wastes” a seat that’s unsold
  - Same for concerts, sports, any event with a fixed date
  - Clothes that are going out of season/fashion
  - Electronics that become obsolete over time
- Any retail setting with inventory constraints
- Often 2 competing effects:
  - The items become less valuable over time, or you have a deadline to sell them
  - You have less stock over time

Simplified example

- You have 1 copy of the item to sell
- There are 2 time periods, today and tomorrow
  - One customer will come in today, a different one tomorrow
- No covariates
- No "discounting" (a dollar tomorrow is as valuable as a dollar today)
- You already have a good estimate of $d(p)$

What price p_1 do you set today? What price p_2 do you set tomorrow?

A couple of observations

If I don’t sell the item today, then *tomorrow* I am solving the same problem that we solved in class last time:

- Maximizing revenue for a single buyer/without capacity considerations
- => The price for tomorrow will be same as simple revenue maximizing price

$$p_2 = \arg\max_p [p \times d(p)]$$

Not true for the price today:

- If I sell the item today, then I lose out on a potential sale tomorrow
- If I don’t sell the item today, I get another chance tomorrow

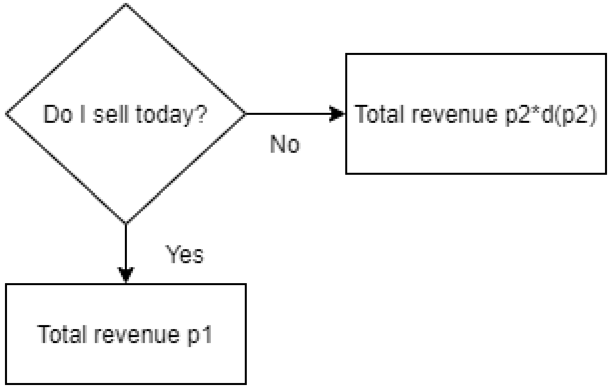
=> I should “take a risk” today to try to sell at a higher price

Solving the example: “Bellman equation”

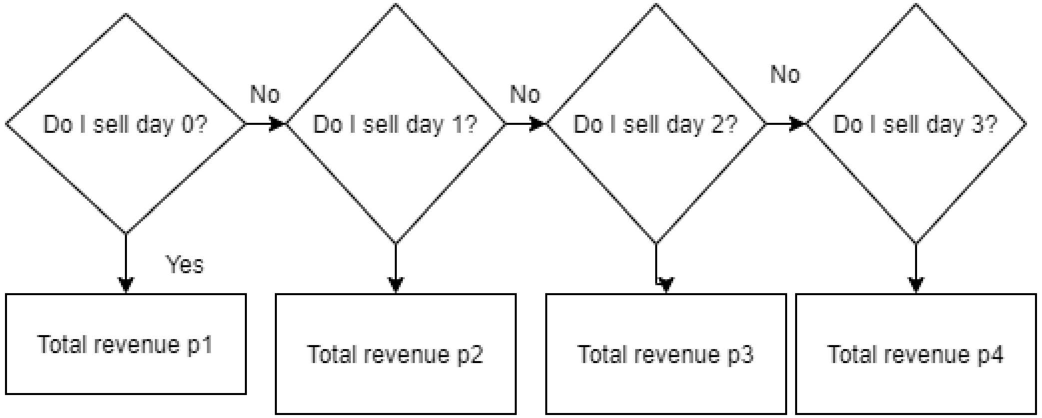
- If I don’t sell today: (happens with probability $1 - d(p_1)$)
  - Then my revenue today is 0
  - Then the expected revenue tomorrow is: $p_2 d(p_2)$
- If I *do* sell today: (happens with probability $d(p_1)$)
  - My revenue today is p_1
  - Then the expected revenue tomorrow is 0
- So, my overall expected revenue is:

$$d(p_1)(p_1 + 0) + (1 - d(p_1))(0 + p_2 d(p_2))$$

- p_2 easy to solve – does not depend on p_1
- *Given* p_2, the above revenue function is only a function of p_1 => *Can optimize* p_1



Bellman equation generally



- You can generalize this idea to selling any number of items sequentially for T days
- Start from Day T: If you still have an item, do single-shot maximization
- Day $\mathrm{T}-1$: Given Day T price, you know expected reward if you still have an item to be sold after Day $\mathrm{T}-1$. And so, you can calculate optimal price for Day $\mathrm{T}-1$.
- Now, you have the expected reward if you still have an item to be sold after Day $\mathrm{T}-2$...

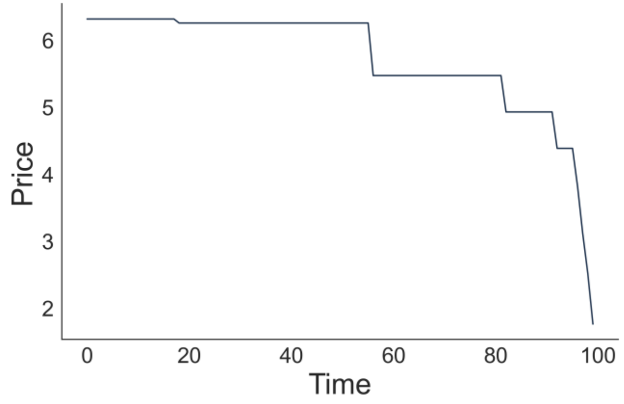
More Bellman equation

- Let $\mathrm{V_t}$ denote: "Expected profit if I still have an item to sell on day t"

$$\mathrm{V_T} = p_T \times d(p_T)$$

$$\mathrm{V_{T-1}} = [p_{T-1} \times d(p_{T-1})] + \left(1 - d(p_{T-1})\right) V_T$$

- Above means: "Value today is revenue today if I sell the item today, or tomorrow's expected revenue if I don't sell the item today"
- For each t, given V_{t+1} we can calculate optimal price p_t
- Keep iterating until you have prices $p_0 \ldots p_T$
- Resulting $\mathrm{V_0}$ is my expected revenue given these prices



Bellman equations: a general idea

- Constructing a tree to reason about what happens tomorrow, and then iterating backwards, is a powerful + flexible algorithmic technique: "dynamic programming"
- Example: What if you have 5 copies of each item?

  Let k denote how many copies of the item I have. Then:

$$\mathrm{V}_{\mathrm{t},0} = 0 \text{ for all } \mathrm{t}$$

$$\mathrm{V}_{\mathrm{t,k}} = d(p_{t,k})\left[p_{t,k} + \mathrm{V}_{\mathrm{t+1,k-1}}\right] + \left(1 - d(p_{t,k})\right) \mathrm{V}_{\mathrm{t+1,k}}$$

  If I sell an item today: Revenue today, plus future revenue from 1 less item

  If I don't sell: Future revenue from same number of items

  Competing effects: Now, less capacity over time → prices should go up (but less time to sell, so prices should go down).

Capacity constraints + over-time pricing in practice

- Dynamic programs/bellman equations are powerful, but often the real world is too complicated
  - Uncertainty in future capacity
  - Future actions of competitors
  - Future demand distributions
  - “Long time horizons” (T is big)
- In theory, dynamic programming can handle the above. In practice, hard to know how to calculate future value.

Approximating dynamic programming

- In the recommendations module, we created “score”(or “index”) functions:
  - Consider future users, through capacity and avg ratings terms in the score function
- With 1 item: V_{t+1} represents my “opportunity cost” if I sell an item today that I could have sold tomorrow.

  Also interpret as “safety net”: if fail to sell the item today, still earn V_{t+1} in expectation
- Instead of doing a full Bellman equation, estimate V_{t+1} through some other means, then plug into the decision problem for today (finding price p_t)
  - Can construct it like we did score functions for recommendations
  - AlphaGo to play Go: V_{t+1} is partially estimated via a neural network

Pricing with capacity summary

- Just like in recommendations, have to think about potential future demand
- Here, potential future demand lets us be “more aggressive” by pricing higher today
- If I can summarize future revenue (V_{t+1}) effectively, then I can optimize today’s prices
- Dynamic programming: start from the end!
- We assumed that customers can’t strategize on when to come – not true!

Selling multiple kinds of items

Price differentiation

Motivation

Motivation 1:

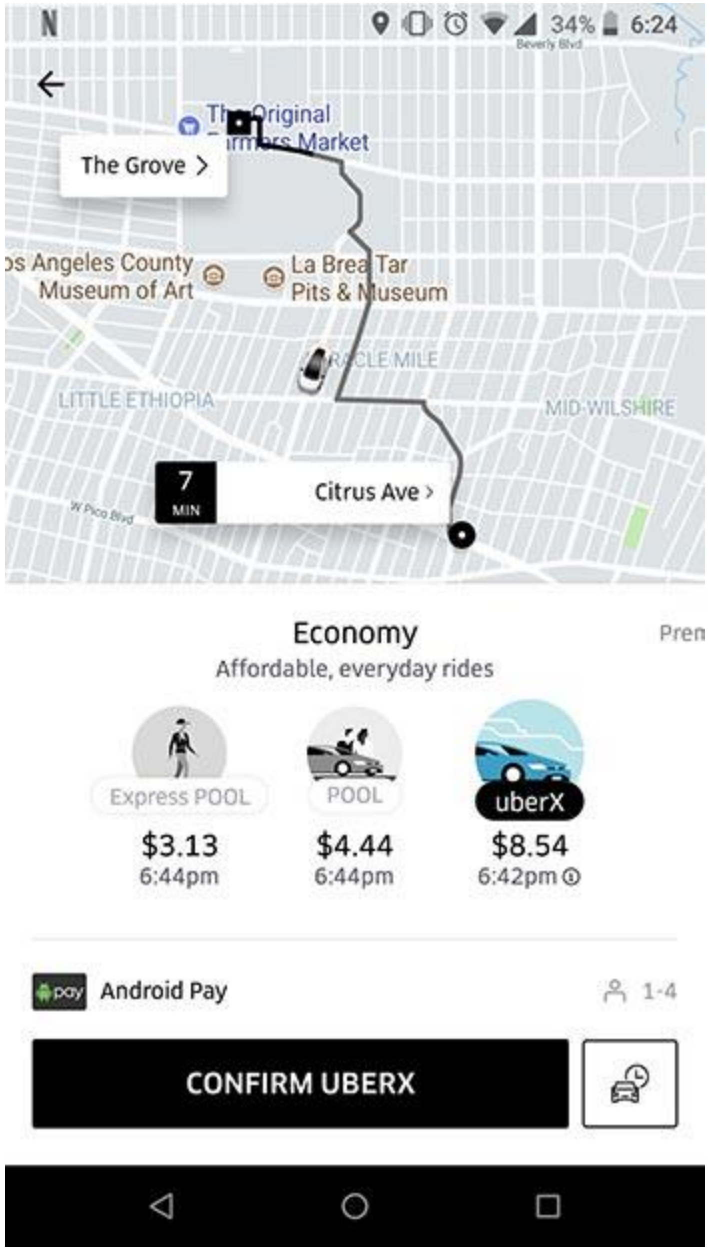
You simply have multiple kinds of products to sell. Different types of clothes, laptops, airline seats, furniture, etc.

Motivation 2:

- Last class: personalized pricing with covariates
- Challenge: Often you can’t (technically, ethically, legally, …) give different prices for the same product to different users based on covariates
- Now: Different “tiers” of service.
  - High quality: First class seats, faster service in Uber/Lyft, luxury goods versions

    Get item “now” versus get item “later”
  - Lower quality: Economy, UberPool/Lyft Wait and Save, …

=> Purposely create tiers of service to earn more money from richer people while earning something from others



Challenges

- Just like pricing over time, now prices for the 2 items depend on each other
  Unlike pricing to different demographic segments without capacity constraints

- Cannibalization: Customers who would have bought the luxury good instead buy the cheaper good because it is available

2-item user behavior model

- Suppose you're selling 2 types of items
- Each person will buy at most one item
  - Each person has a *private valuation* v_1 for item 1 and v_2 for item 2
- Suppose you offer the items at price p_1 and p_2, respectively
- How does the person make their decision?

  Utility from item j at price p_j is $v_j - p_j$
- Person i buys

  Neither item if $v_1 < p_1$ and $v_2 < p_2$

  Item 1 if $v_1 \geq p_1$ and $v_1 - p_1 \geq v_2 - p_2$

  Item 2 if $v_2 \geq p_2$ and $v_2 - p_2 \geq v_1 - p_1$

In more detail

How does the person make their decision? Person i buys

- Neither item if $v_1 < p_1$ and $v_2 < p_2$
- Item 1 if $v_1 \geq p_1$ and $v_1 - p_1 \geq v_2 - p_2$
- Item 2 if $v_2 \geq p_2$ and $v_2 - p_2 \geq v_1 - p_1$

Revenue in 2 item model

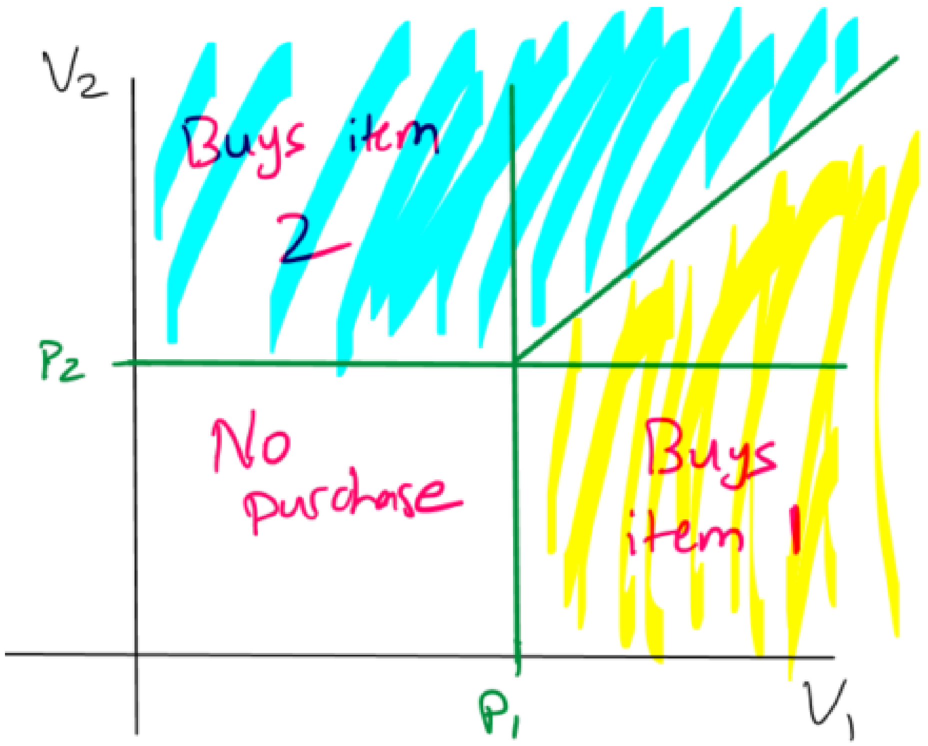
For a set of prices (p_1, p_2), let

- $\mathrm{d}_1(p_1, p_2)$ be fraction of people who buy item 1 (Yellow Region)
- $\mathrm{d}_2(p_1, p_2)$ be fraction of people who buy item 2 (Blue Region)

- Then, revenue is:

$$p_1 \times \mathrm{d}_1(p_1, p_2) + p_2 \times \mathrm{d}_2(p_1, p_2)$$

- Given functions $\mathrm{d}_1, \mathrm{d}_2$, can solve for optimal prices



Cannibalization

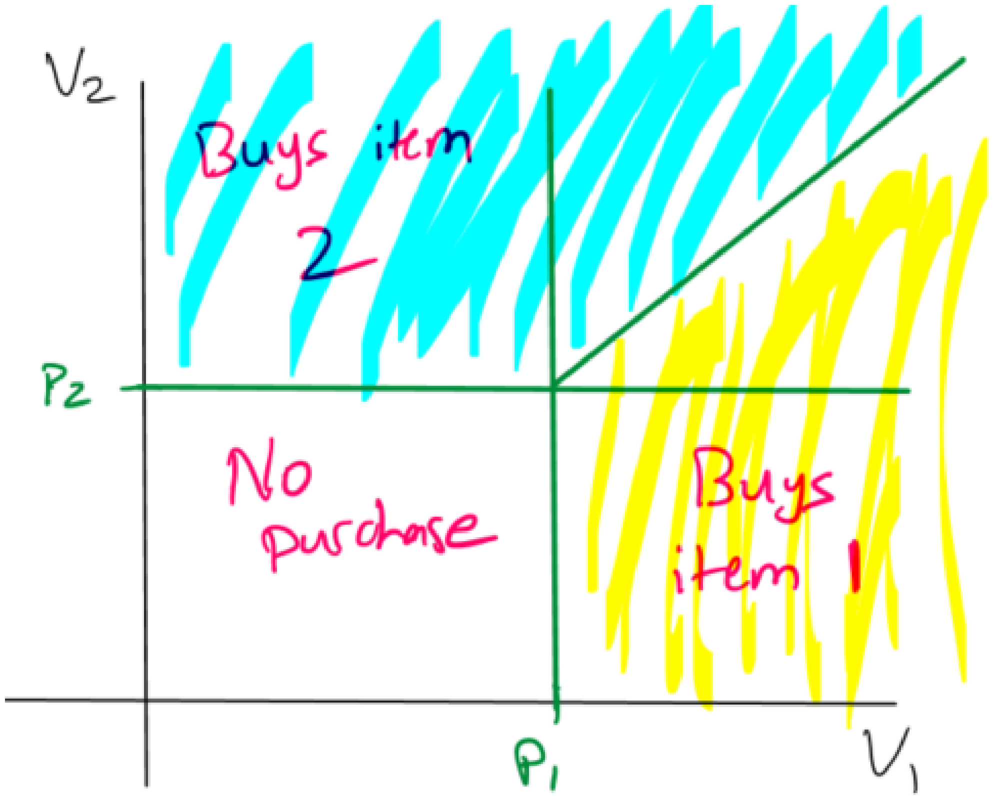
Now, each price affects other item.

Suppose you decrease p_1 (make item 1 cheaper)

Then:

- Earn less money in yellow region ↓
- Yellow region becomes bigger ↑

  White region becomes smaller
- **Blue region becomes smaller ↓**



Demand estimation with multiple items

- With a single item, we suggested machine learning approach to estimate: $\mathrm{d}(\mathrm{p},\mathrm{x}) \stackrel{\text{def}}{=} 1 - F_{p|X}\,(p \mid X = x)$
- Assume we have user i with covariates x_i
- Now, would need to estimate $\mathrm{d}_1(p_1, p_2, x_i)$ and $\mathrm{d}_2(p_1, p_2, x_i)$
- Gets very hard, very quickly
- Approach 1: Use a *multi-class* classification algorithm $g(p_1, p_2, x_i)$
  [Buy nothing, buy item 1, buy item 2] and then extract class probabilities
  (sci-kit learn: use **predict_proba** with any multi-class classifier)
- Approach 2: (From beginning of today's class)
  - Use user and item vectors, i.e., $(\mathrm{p}_1, \mathrm{p}_2, u_i \cdot w_{\text{item }1}, u_i \cdot w_{\text{item }2})$

Substitutes and complements

- So far: motivation -- we have multiple products to sell, that appeal to different customers

  “cheaper” and “more expensive” product
- Items are “substitutes”: people only buy at most one kind of item
- Sometimes, items are “complements” – buying one item makes the other item more attractive
  - Soda + popcorn at movie theater
  - iPhone and Macbook and Apple Watch and Apple TV and …
- Then, reducing one item’s price might induce you to buy more overall
  - An item is a “loss leader”

Summary and next time

- So far: Demand estimation and revenue-maximizing prices
    - Capacity constraints
    - One or two items
    - Potentially allowed to give different prices to different users
    - Set different prices over time
- Simplifying assumptions:
    - No competition from other sellers
    - No over-time dynamics
- Next time
    - Briefly talk about the above
    - Big picture: putting things together
        - Details of class competition
        - Pricing in ride-hailing and online marketplaces

Questions?