ISOM 2700: Operations Management

Session 14. Demand Forecasting

Yiwen Shen
Dept. of ISOM, HKUST
Fall 2025

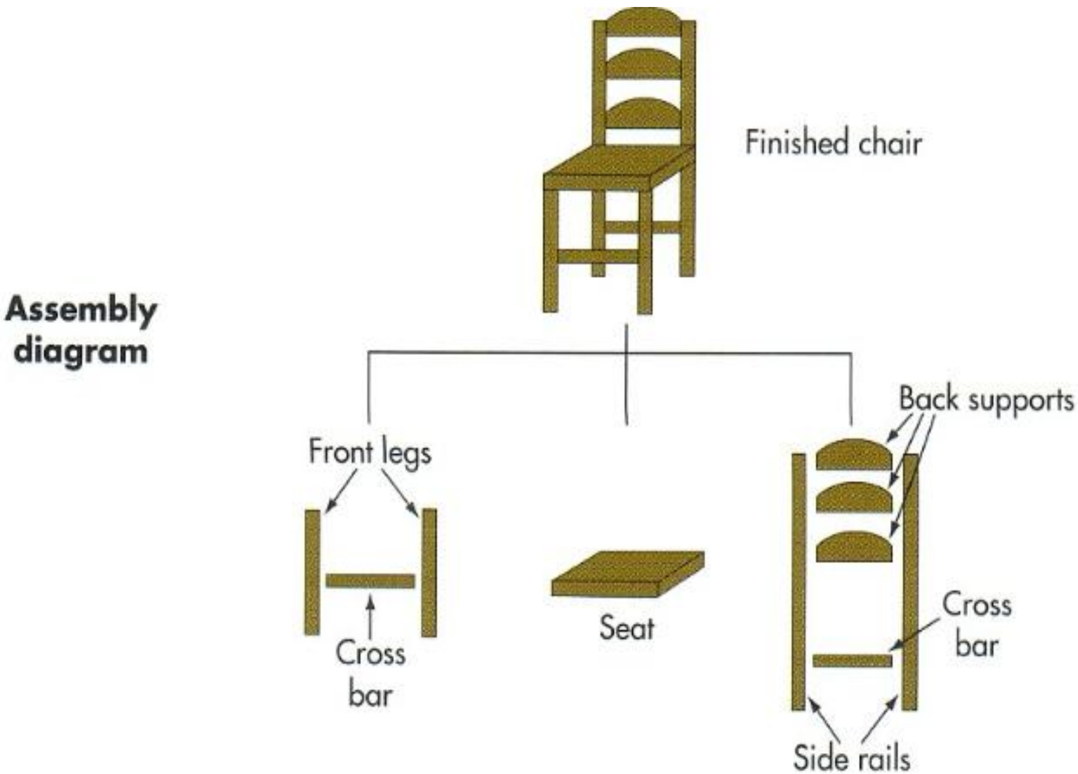
Agenda

- **Overview of forecasting**
- Time-series analysis forecasting
- Forecasting accuracy measures

Demand forecasting is essential in OM

- Capacity planning, inventory management, revenue optimization, supply chain management
  - Demand forecasting is an indispensable input
  - Decision tree, linear programming, queueing analysis, etc
- Examples:
  - Hospital: incoming patients and their illness severity
  - Airline and hotel: passengers and their route/stay
  - Queueing: arrival rate and service time needed
  - University: students admitted & career opportunities

Independent and Dependent Demand



Independent Demand

- Based on final customer demand
- Cannot be derived directly from other products/services

Dependent Demand

- Coming from demand for another product/service
- Can be derived from other products/services

We focus on the forecasting of independent demand, which is more important

Types of Forecasting

- Strategic forecasts: mid- and long-term
  - Strategic decisions (enter or exit a market), capacity-planning (building a new factory)
- Tactical forecasts: short-term
  - Day-to-day level, estimating waiting time and demands, scheduling and staffing decisions
- Qualitative forecasting:
  - Subjective, flexible, forward thinking, easy to incorporate personal judgements and thoughts
- Quantitative forecasting:
  - Objective, data-driven, efficient and automated

Principles of Demand Forecasting

- Forecasting is only forecasting: Never perfect
- The longer the forecast horizon, the worse the forecast
  - Because we have more uncertainty
- Aggregate forecasts are usually more accurate
  - Incorporate more factors to mitigate noise
- Good forecasts need to balance business acumen and quantitative analysis

Qualitative Forecasting Methods

- Not required
- Market research
  - Collect data in a variety of ways (surveys, interviews, etc.) to test hypothesis about the market
- Forecast combination:
  - Average the forecasts from multiple participants or models
- Consensus building:
  - Let participants discuss and explain their forecasts via meetings
- Prediction markets:
  - Process forecasting information from the market (e.g., prices, betting odds)
- Need to be aware of potential behavioral bias in qualitative forecasting (e.g., overconfidence, anchoring, group thinking)

Agenda

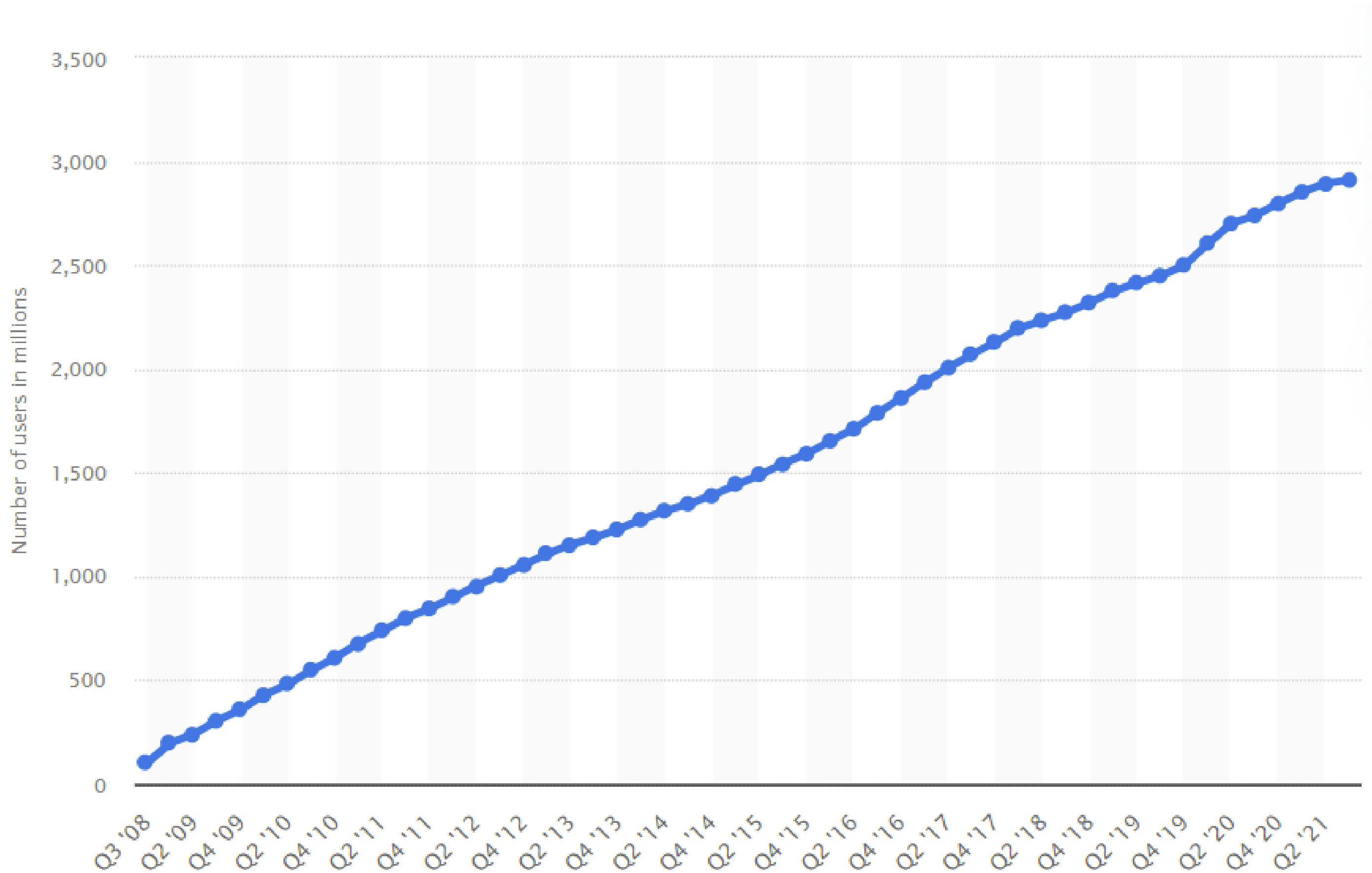
- Overview of forecasting
- **Time-series analysis forecasting**
- Forecasting accuracy measures

Time Series Analysis

- Time Series Analysis is commonly used for quantitative forecasting
- Based on only old demand (or forecasts) data; a form of extrapolation
- Underlying assumptions:
  - Data in the past provides useful information for the future
  - Past patterns (trend, seasonality) will prevail in the future
- Cannot effectively respond to new, emerging trends or breakthroughs

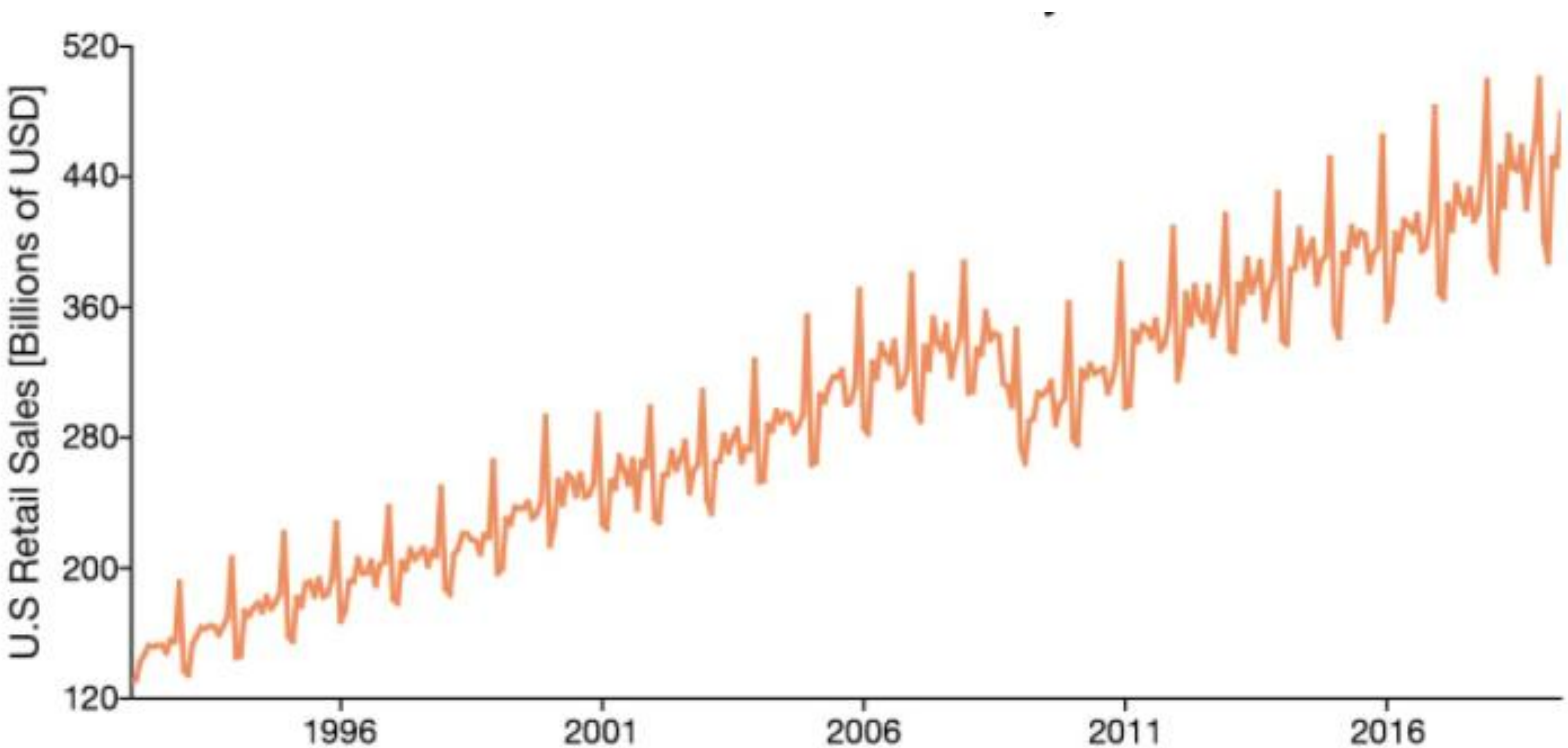
Components of Time Series: Trend

Number of Facebook active users grows over time



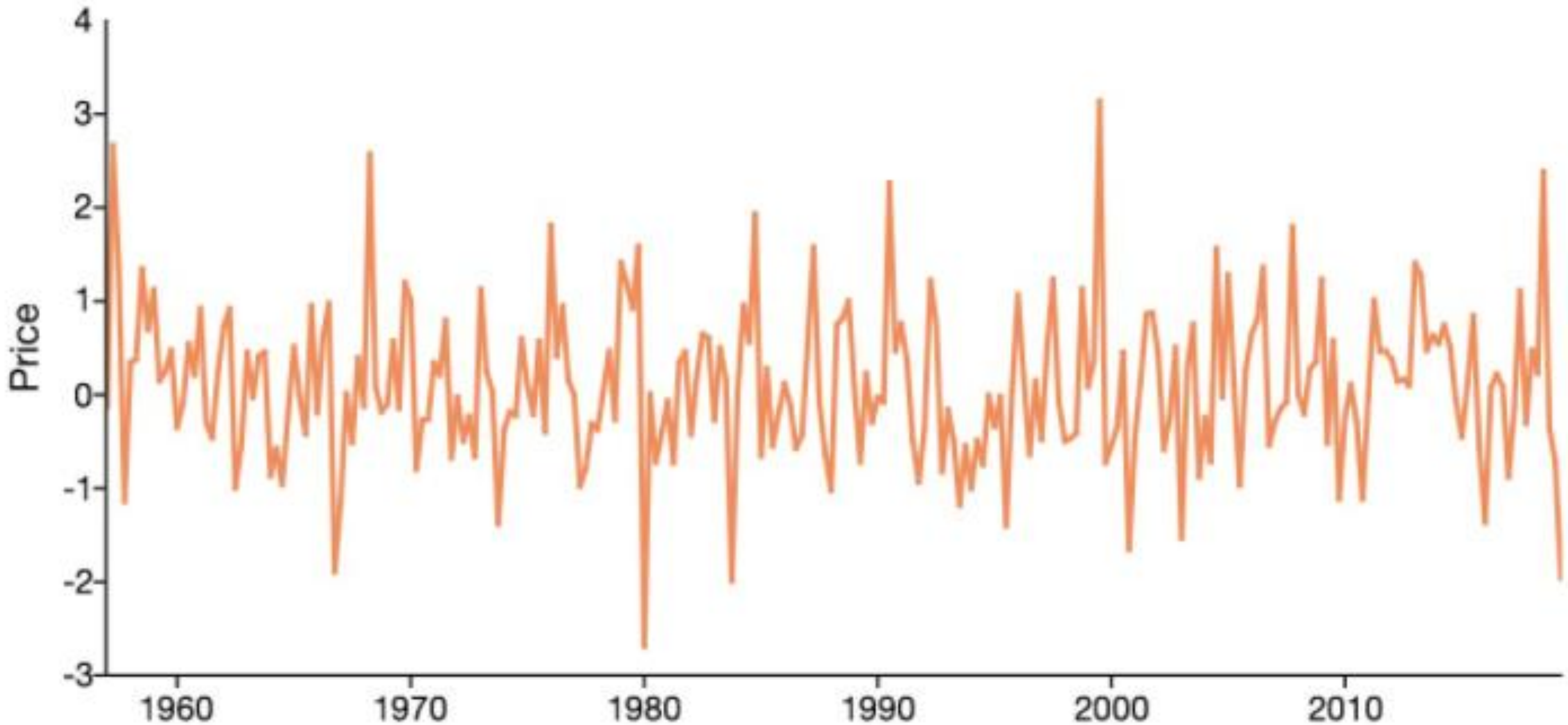
Components of Time Series: Seasonality

US retail sales is usually higher at the end of year



Components of Time Series: Random Variation

Variability over time



Time-series Analysis Forecasting

- Modeling trends
  - e.g., double exponential smoothing
- Modeling seasonality
  - e.g., seasonality index
- Control/reduce random noise
- Other issues: autocorrelation, structural break, etc..

Trade-off in Forecasting Models

- Responsiveness: we want the forecasts to be able to capture recent changes in the demands
  - e.g., if there is recently more demand for the product
  - this means we should rely more on the "recent" demand data

- Robust/Stability: we want the forecasts to be robust to random noises in realized demands
  - i.e., the forecasts should not be too volatile
  - this means we should smooth over the realized demands over more periods in the past

- Thus, there is a trade-off between responsiveness and robust/stability

Forecasting Methods

- Naïve model: $F_t = A_{t-1}$
  - F_t is the forecast for period *t*, A_{t-1} is the realized demand in period $t-1$
  - Highly sensitive to noise, largely incorrect and unstable

- Simple moving average and weighted moving average

- Exponential smoothing

Simple Moving Average

Week	Demand
1	650
2	678
3	720
4	785
5	859
6	920
7	850
8	758
9	892
10	920
11	789
12	844

$$F_t = \frac{A_{t-1} + A_{t-2} + A_{t-3} + \ldots + A_{t-n}}{n}$$

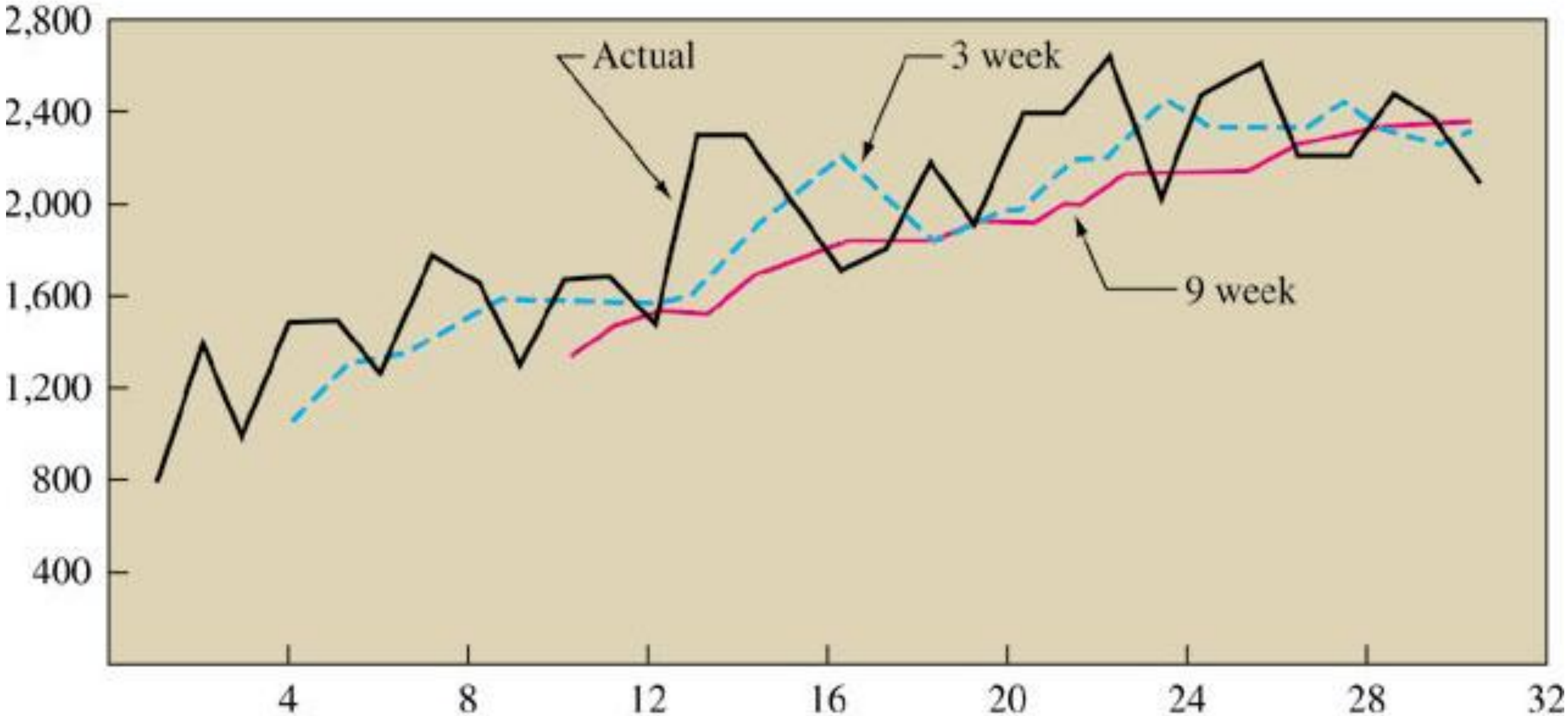
- Take the average of the realized demands in past *n* periods
- Parameter *n* controls the length of look-back window

Effect of Look-back Window

$$F_t = \frac{A_{t-1} + A_{t-2} + A_{t-3} + \ldots + A_{t-n}}{n}$$

- What is the role of *n* (look-back window length) here?
- Larger n: we average over a longer past period
  - Pros: smooth out the noise, make the forecasts more robust
  - Cons: less responsive to most recent demand innovation
- Trade-off between responsiveness versus robustness/stability
- We need to choose n depending on the product feature and forecasting performance metrics

Effect of Look-back Window: Example



The forecasts with a longer window (9 weeks) is smoother, but responds more slowly to actual demand

Weighted Moving Average

- Weighted moving average: allowing different weights to be placed on each past observation

- Model:

$$F_t = w_1 A_{t-1} + w_2 A_{t-2} + \dots + w_n A_{t-n}$$

$$\sum_{i=1}^{n} w_i = 1$$

- Parameters: selected by trial and error
  - Look-back window length n
  - Weights: w_i for $i = 1,2,\dots,n$

Weighted Moving Average

- Model:

$$F_t = w_1 A_{t-1} + w_2 A_{t-2} + \ldots + w_n A_{t-n}$$

$$\sum_{i=1}^{n} w_i = 1$$

- If all $w_i = 1/n$, we get back to the simple moving average
- Improve responsiveness: Allocate more weight to recent data, i.e., bigger w_i for smaller i
- Again, this will make forecasts more sensitive to noise: trade-off between robustness and responsiveness

Weighted Moving Average: Example

Question: Given the weekly demand and weights, what is the forecast for the 4th period or Week 4?

Week	Demand
1	650
2	678
3	720
4	693.4

Weights:

t-1	0.5
t-2	0.3
t-3	0.2

$693.4 = 720\times0.5+678\times0.3+650\times0.2$

Note that the weights place more emphasis on the most recent data, that is time period "t-1"

Simple vs. Weighted Moving Average

- Simple Moving Average
  - Based on average past demand
  - Equal importance to each observation
- Weighted Moving Average
  - More recent data are usually given more significance than older data
- Drawbacks for both:
  - Need to continually carry large historical data (e.g., consider using 60-day moving average for each of 100,000 items)
  - Data before *n* periods are entirely dropped

Exponential Smoothing

- Model:

$$F_t = \alpha A_{t-1} + (1-\alpha)F_{t-1}$$
$$= F_{t-1} + \alpha(A_{t-1} - F_{t-1})$$

Where :

F_t = Demand Forcast for the coming time period

F_{t-1} = Demand forecast in the past time period

A_{t-1} = Actual demand in the past time period

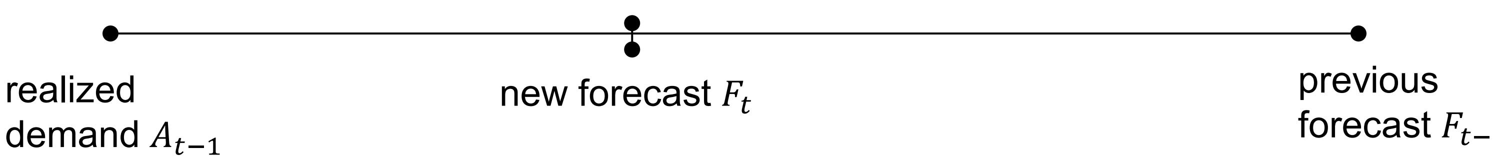
α = Alpha smoothing constant

- Input: most recent forecast and demand
- Parameter: a smoothing constant α (between 0 and 1)

Exponential Smoothing

- Model:

$$F_t = \alpha A_{t-1} + (1-\alpha)F_{t-1}$$
$$= F_{t-1} + \alpha(A_{t-1} - F_{t-1})$$



- The new forecast is a weighted average of previous forecast and realized demand

Exponential Smoothing: Example

- Assume that the last month's forecast was 1050 units
- It turns out that 1000 units were actually demanded
- Assume smoothing constant $\alpha = 0.20$
- Forecast for this month:

$$1050 + 0.20 \times (1000-1050) = 1040$$

- Intuition: New forecast revises the old forecast in the direction of actual demand

Smoothing Constant α: Trade-off

- α controls by how much we revise the previous forecast in the direction of realized demand

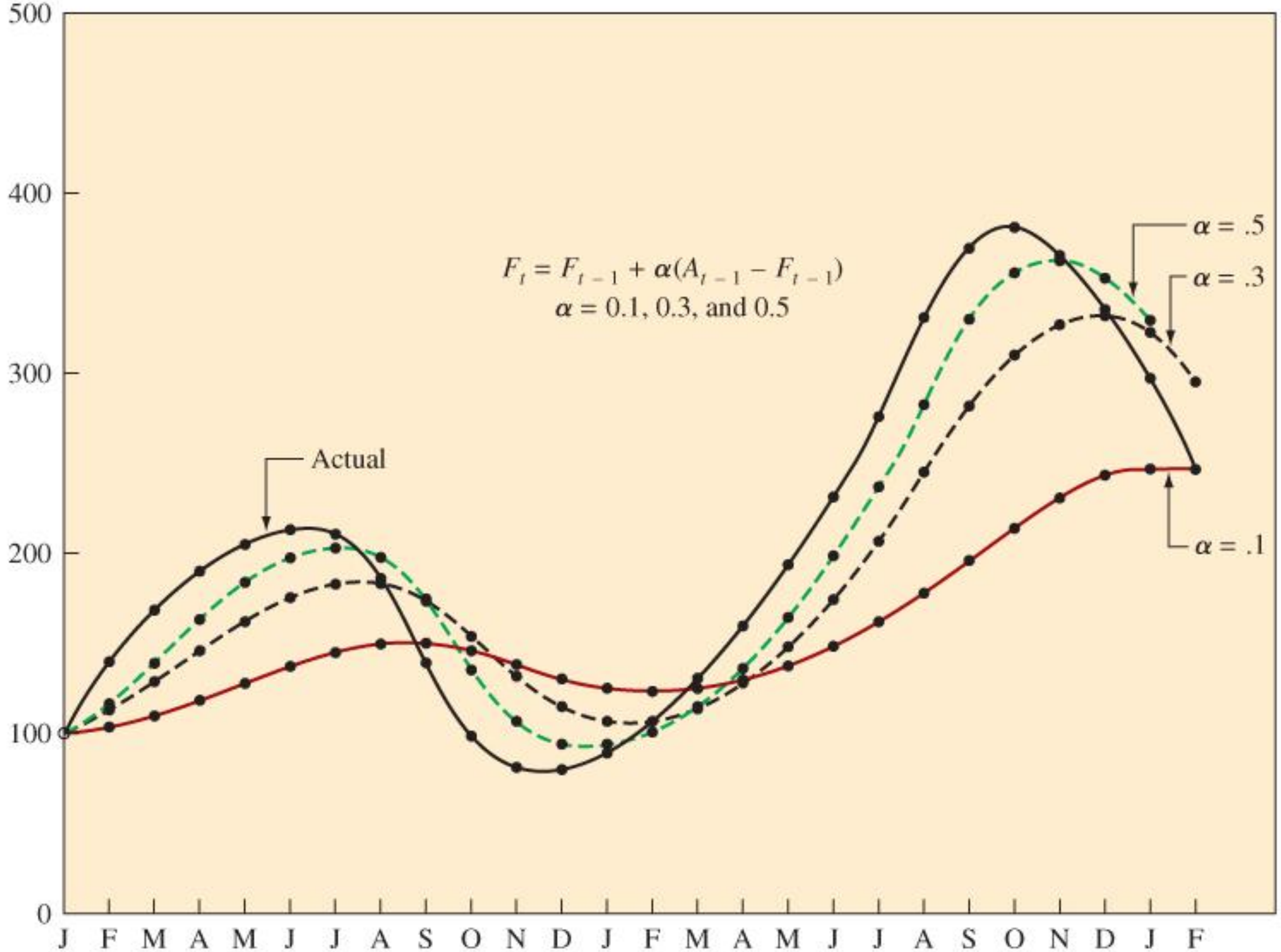
$$F_t = F_{t-1} + \alpha(A_{t-1} - F_{t-1})$$

- Large α means we rely more on the most recent demand:
  - Forecasts are more responsive to new demand innovation
  - But more sensitive to random noise , less robust and stable

- Small α means we rely more on original forecast:
  - Forecasts are more robust and less sensitive to random noise
  - But less responsiveness to recent data

- We choose α based on the importance of new information and forecast accuracy metrics

Effect of Smoothing Constant α



Agenda

- Overview of forecasting
- Time-series analysis forecasting
- **Forecasting accuracy measures**

Measuring Forecast Accuracy

You are a marketing analyst for McDonalds and have the following sales forecasts ($M) using two methods.

Month	Actual Demand	Method 1 Forecast	Method 2 Forecast
1	100	60	100
2	100	130	100
3	200	200	150
4	200	270	200
5	400	340	250

Measuring Forecast Accuracy

- Suppose we have the actual demand and forecast for period $t = 1,2, \ldots, T$
- Forecast error is the difference between the actual value and the predicted value

$$Error_t = Actual_t - Forecast_t$$

- Mean absolute deviation (MAD)

$$MAD_t = \frac{\sum_{i=1}^{t} |Actual_i - Forecast_i|}{t}$$

- Tracking signal: ratio of cumulative error and MAD

$$TS_t = \frac{\sum_{i=1}^{t} (Actual_i - Forecast_i)}{MAD_t}$$

Calculation of MAD and TS

- Mean absolute deviation (MAD)

$$MAD_t = \frac{\sum_{i=1}^{t} |Actual_i - Forecast_i|}{t}$$

- Note that we take absolute value of forecasts error in the numerator : both overshooting and undershooting matter

- Tracking signal: ratio of cumulative error and MAD

$$TS_t = \frac{\sum_{i=1}^{t} (Actual_i - Forecast_i)}{MAD_t}$$

- We calculate the cumulative error in the numerator: captures trend in the cumulative errors
- The denominator is MAD, which must be calculated first

Measuring Forecast Accuracy

- Mean absolute deviation (MAD):
  - Accounts for both overshooting and undershooting of the realized demands
  - Measures the average of deviation from true value
- Tracking signal (TS):
  - Accounts for cumulative errors over time
  - Warns when there are unexpected departures of realized demands from the forecasts (same idea as control charts)
  - A general criteria: $|TS| > 3.75$ implies the forecast is poor
- They are two different accuracy measures; both can be used for comparing forecasting models

Measuring Forecast Accuracy

- MAD and TS are two different accuracy measures; both can be used for comparing forecasting models
- Better performance in one measure does NOT imply a better performance in the other
- For example, a model can have a smaller MAD but larger deviation by TS, and vice versa

TS and MAD: Examples

- Assume the demand is constant 100 every period
- Consider a method that gives forecast at 95 every period
  - MAD equals 5 every period (fine)
  - But TS keeps increasing to high levels (poor)
- Consider a method that gives forecast 150 and 50 alternatively
  - MAD equals 50 every period (poor)
  - But TS takes value of 0 and 1 alternatively (fine)

Method 1: Forecast Accuracy

Month	Actual	Forecast	Error	∑|Error|	MAD	∑(Error)	TS
1	100	60					
2	100	130					
3	200	200					
4	200	270					
5	400	340					

Error = Actual – Forecast   MAD = ∑|Error| / n   TS = ∑(Error) / MAD

Method 1: Forecast Accuracy

Month	Actual	Forecast	Error	∑|Error|	MAD	∑(Error)	TS
1	100	60	**40**				
2	100	130	**-30**				
3	200	200	**0**				
4	200	270	**-70**				
5	400	340	**60**				

Error = Actual – Forecast

MAD = ∑|Error| / n

TS = ∑(Error) / MAD

Method 1: Forecast Accuracy

Month	Actual	Forecast	Error	∑|Error|	MAD	∑(Error)	TS
1	100	60	40	**40**		**40**	
2	100	130	-30	**70**		**10**	
3	200	200	0	**70**		**10**	
4	200	270	-70	**140**		**-60**	
5	400	340	60	**200**		**0**	

Error = Actual – Forecast

MAD = ∑|Error| / n

TS = ∑(Error) / MAD

Method 1: Forecast Accuracy

Month	Actual	Forecast	Error	∑|Error|	MAD	∑(Error)	TS
1	100	60	40	40	**40**	40	
2	100	130	-30	70	**35**	10	
3	200	200	0	70	**23**	10	
4	200	270	-70	140	**35**	-60	
5	400	340	60	200	**40**	0	

Error = Actual – Forecast

MAD = ∑|Error| / n

TS = ∑(Error) / MAD

Method 1: Forecast Accuracy

Month	Actual	Forecast	Error	∑|Error|	MAD	∑(Error)	TS
1	100	60	40	40	40	40	**1.00**
2	100	130	-30	70	35	10	**0.29**
3	200	200	0	70	23	10	**0.43**
4	200	270	-70	140	35	-60	**-1.71**
5	400	340	60	200	40	0	**0.00**

Error = Actual – Forecast

MAD = ∑|Error| / n

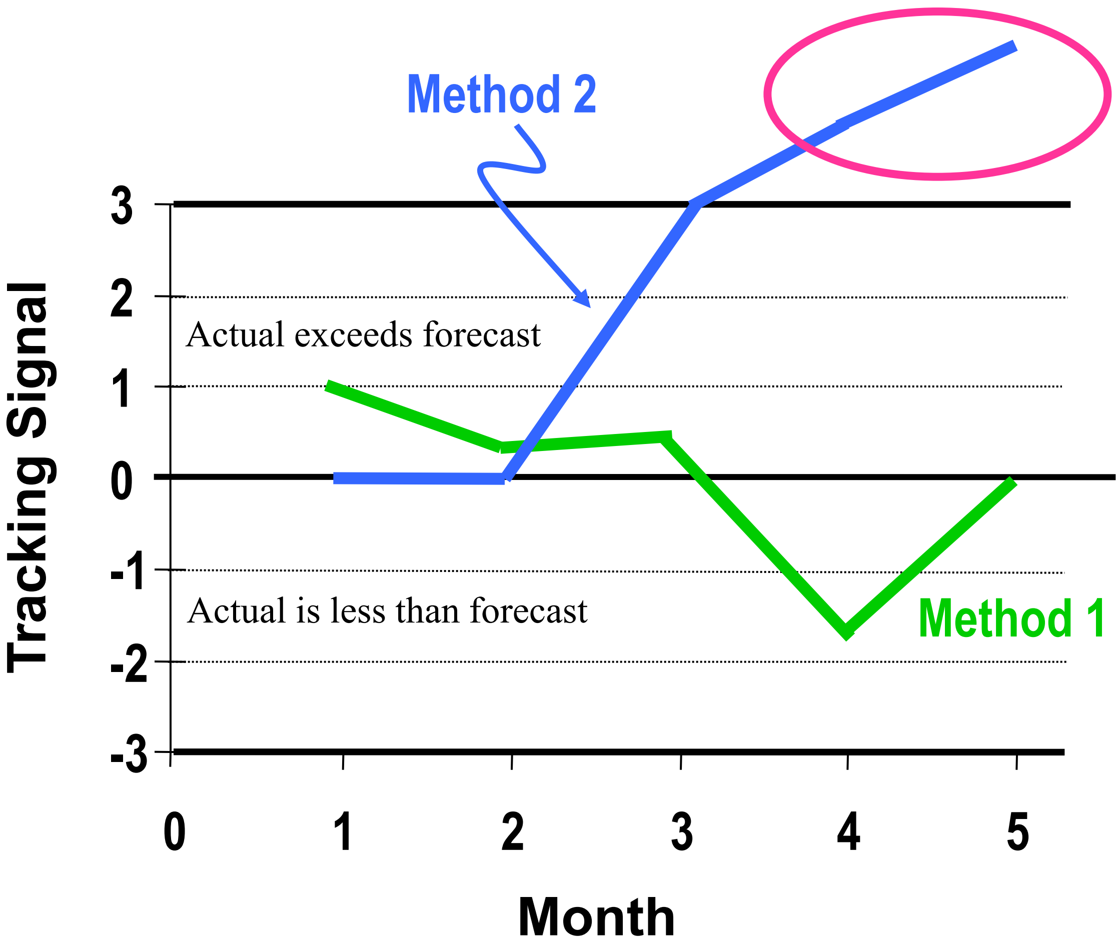
TS = ∑(Error) / MAD

Method 2: Forecast Accuracy

Month	Actual	Forecast	Error	∑|Error|	MAD	∑(Error)	TS
1	100	100	0	0	0	0	**0.00**
2	100	100	0	0	0	0	**0.00**
3	200	150	50	50	16.7	50	**3.00**
4	200	200	0	50	12.5	50	**4.00**
5	400	250	150	200	40	200	**5.00**

Error = Actual – Forecast   MAD = ∑|Error| / n   TS = ∑(Error) / MAD

Interpreting Tracking Signals



Same MAD (40), but TS shows a deviation for Method 2

ISOM 2700: Operations Management

Session 15. Inventory Management I: EOQ Model

Yiwen Shen
Dept. of ISOM, HKUST
Fall 2025

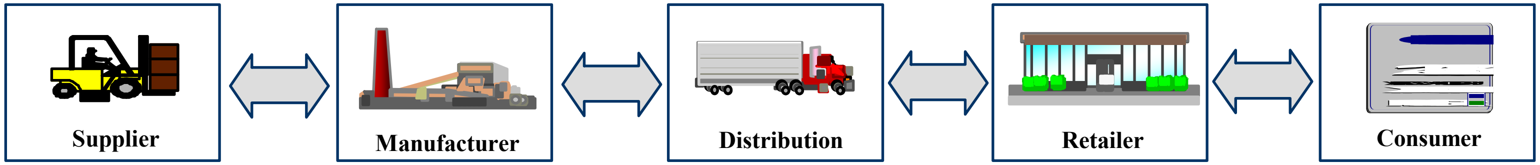
Agenda

- **Overview of inventory management**
- EOQ model: Solution and analysis
- Newsvendor model: Solution and extensions

What is Inventory

- Inventory is a stock of goods awaiting consumption
  – This can happen in each step during the production process

- Examples of inventory:
  – Raw materials
  – Work in process
  – Supplies
  – Finished goods



Inventory Management

- Goal: Matching (inventory) supply with (customer) demand

- Inventory management is crucial:
  - Inventory can be a firm's largest and most important asset
  - Convergence point of supply chain and touchpoint of customer
  - Multiple stakeholders involved: manufacturing, purchasing, logistics, marketing, accounting, etc

- Key components:
  - Inventory tracking, order management, transfer management, reporting and analysis, purchasing, shipping and delivery

- Cutting-edge development:
  - Artificial intelligence, internet of things, block chain, big data analytics

Challenges of Inventory Management

- High costs
  - Ordering cost (per order)
  - Holding cost (per unit of inventory per unit time)
    - Spoilage, obsolescence, capital, financing, etc.
  - Stock-out or shortage cost (per unit of lost sales)
    - Occurs when the demand for inventory exceed its supply
    - Customer's goodwill cost, delivery cost or penalty cost

- Difficult to control
  - Uncertainty in the demand
  - Random lead-time for delivering
  - Quality variance

Target Inventory Problem after Covid-19

Video

Inventory Types (I)

- Pipeline Inventory
  - The time a flow unit needs to spend in the process from input to output (Recall Little's Law)
- Seasonal Inventory
  - Occurs when capacity is fixed and demand is variable
  - E.g., produce more Christmas trees in October for sales in December
- Cycle Inventory
  - Economy of scale: beneficial to produce many units in batches
  - This leads to cycle inventory during the production process

Inventory Types (II)

- Inventory as Buffers
  - Use inventory as buffers between activities
  - Enhance the independence and robustness of production process

- Safety Inventory
  - To hedge the unpredictable variation from stochastic demand
  - Trade-off: order too many versus too few

Inventory Models

- Key question: how to find optimal inventory quantity?
- Economic Order Quantity (EOQ) Model
  - Deterministic demand and long lifecycle products: inventory can be held for a long period without depreciation
- Newsvendor Model
  - Uncertainty demand and short lifecycle products: inventory becomes valueless after a certain period
- Fixed-time period model: (not required)
  - Usually used when a group of items is ordered together

Agenda

- Overview of inventory management
- **EOQ model: Solution and analysis**
- Newsvendor model: Solution and extensions

Example: Computer Retailer

- M&S has a stable demand for a line of computer it offers. Each week there is a demand for 100 computers. M&S incurs a fixed cost of \$5000 every time it places an order. The marginal cost of a computer is \$400, and the shop's holding cost of inventory is approximately 1% per week. Assume orders are received one week (lead time) after ordering.

- Question:
  - How often should M&S order?
  - How many units should be ordered each time?

Economic Order Quantity Model

- This types of inventory problem is solved by Economic Order Quantity (EOQ) model

- Decisions:
  - How often should an order be placed?
  - How many units should be ordered each time?

- Key structure (assumptions):
  - A constant demand rate for inventory (no uncertainty)
  - Replenishment occurs a fixed time (lead time) after ordering
  - A fixed cost for each time we order
  - A constant holding cost per unit of inventory

Inventory Level Pattern

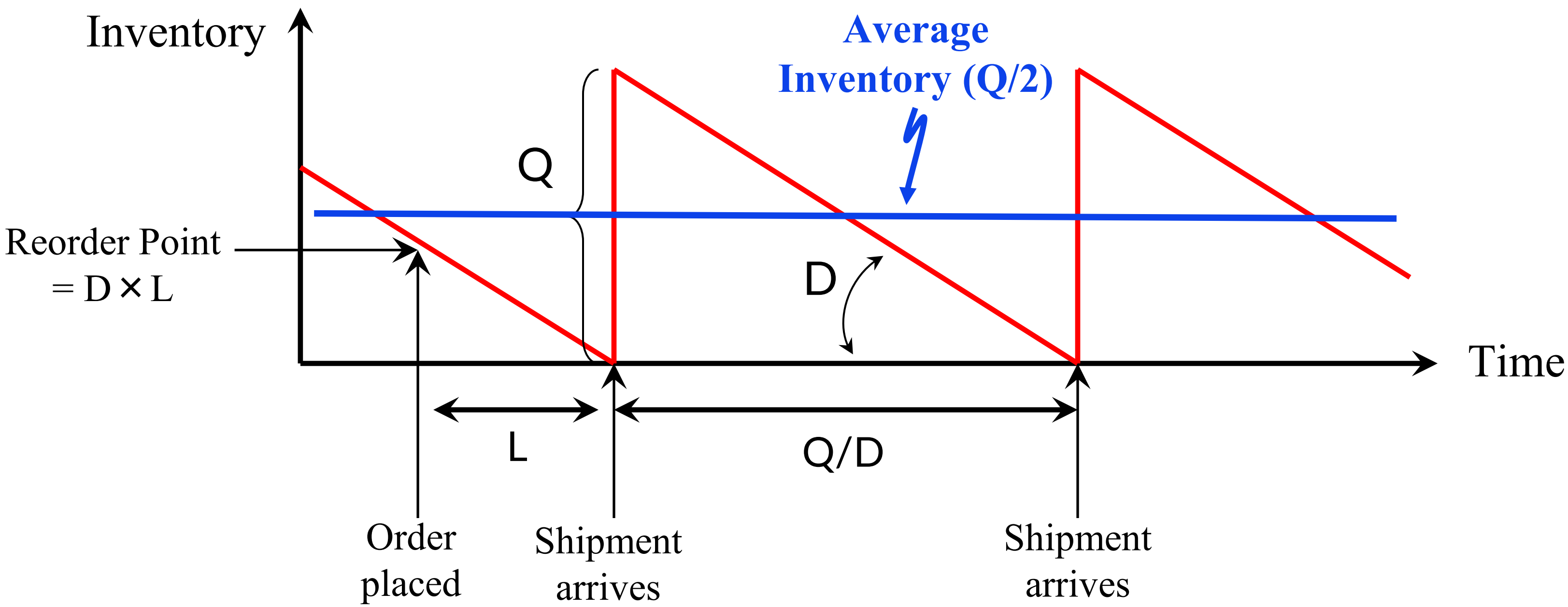
- We first consider how the firm's inventory level changes with time

- Given constant demand rate, the inventory will drop linearly with time due to demand when no new product is received

- When a new batch of products is ordered, it will arrive one week later

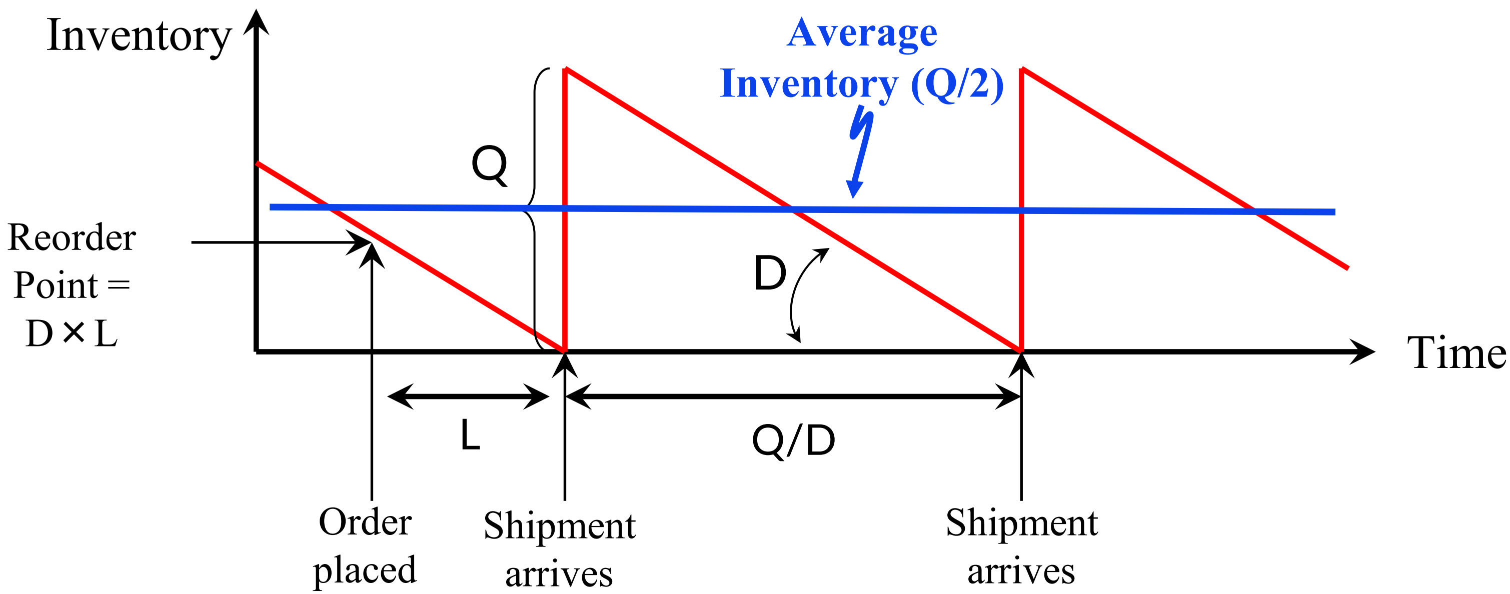
- When a new batch of products is received, your inventory level has an upward jump

Inventory "Saw-Tooth" Pattern



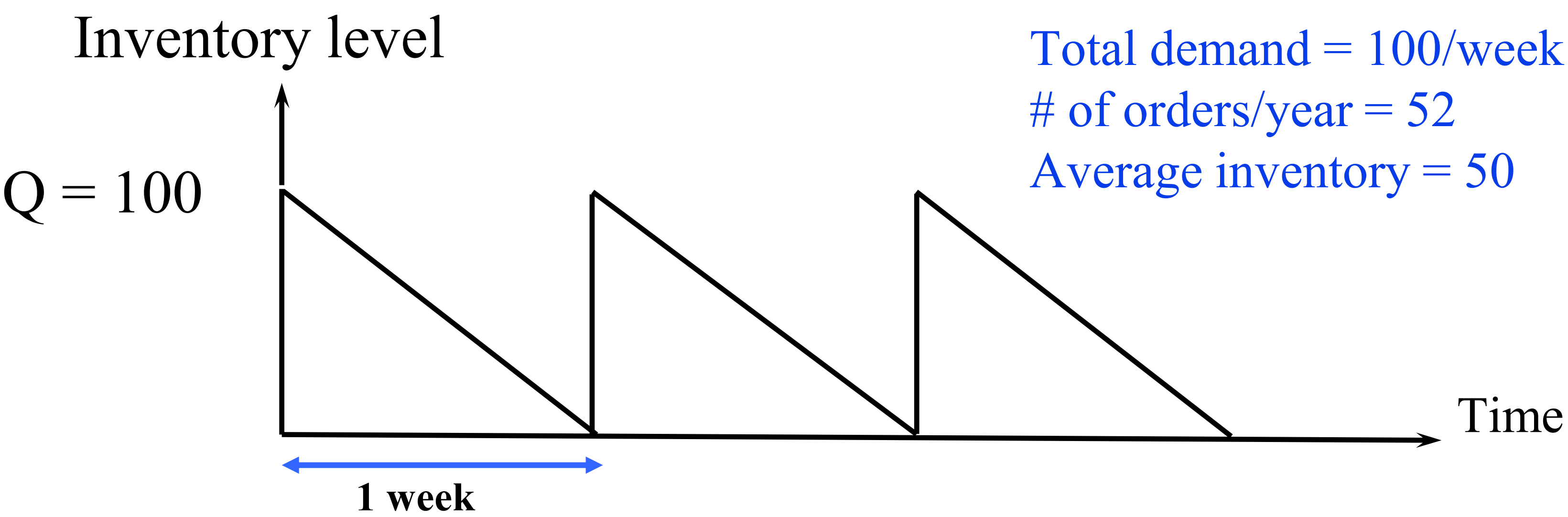
- D: demand rate, i.e., demand quantity per unit of time
- Q: order quantity each time we place the order (what we need to choose)
- L: lead time, i.e., how long you wait before you get the inventory after you place the order

Inventory "Saw-Tooth" Pattern

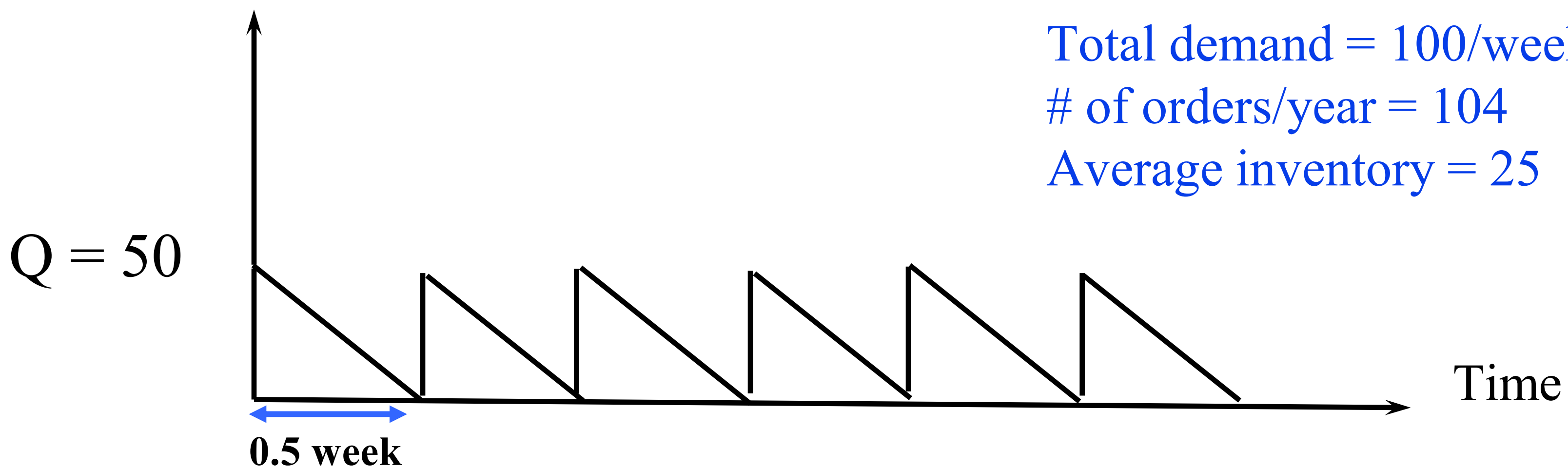


- As the demand rate is deterministic, we always order such that the new orders arrive when the inventory level drop to zero
  - This reduces holding cost, leading to a zig-zag inventory pattern
- Q/D: Time between shipments; D/Q: order frequency each unit of time
- Reorder point: the inventory level when we start to place a new order, given by D ×L

Order Frequency



Total demand = 100/week
# of orders/year = 52
Average inventory = 50



Total demand = 100/week
# of orders/year = 104
Average inventory = 25

Decision: Order Quantity

- The more frequent we order, the fewer inventory we need to order each time

- This decreases the average holding cost (proportional to inventory level)
- …but increases the order (set-up) cost as we need to order for more times

- Note that the lead time does NOT affect the total cost; it only affects when we place the order

Economic Order Quantity (EOQ) Model: Total Cost

Given (time unit should be same for all quantities):

D - Demand per unit of time (year, week, …)

S - Setup or Order Cost ($/setup; $/order)

H - Marginal holding cost ($/per unit inventory per unit of time)

Given we order Q each time,

- Each batch of inventory can sustain the demand for Q/D units of time (this is the time gap between ordering)
- In a unit of time, we need to order in total D/Q batches
- The average inventory level is Q/2

Economic Order Quantity (EOQ) Model: Total Cost

Given (time unit should be same for all quantities):

D - Demand per unit of time (year, week, …)

S - Setup or Order Cost ($/setup; $/order)

H - Marginal holding cost ($/per unit inventory per unit of time)

Given we order Q each time, calculate the total inventory cost in per unit of time

- Holding cost = $H \times Q/2$
- Ordering cost = $S \times D/Q$
- Total cost = $H \times Q/2 + S \times D/Q$

EOQ Model: Optimal Q

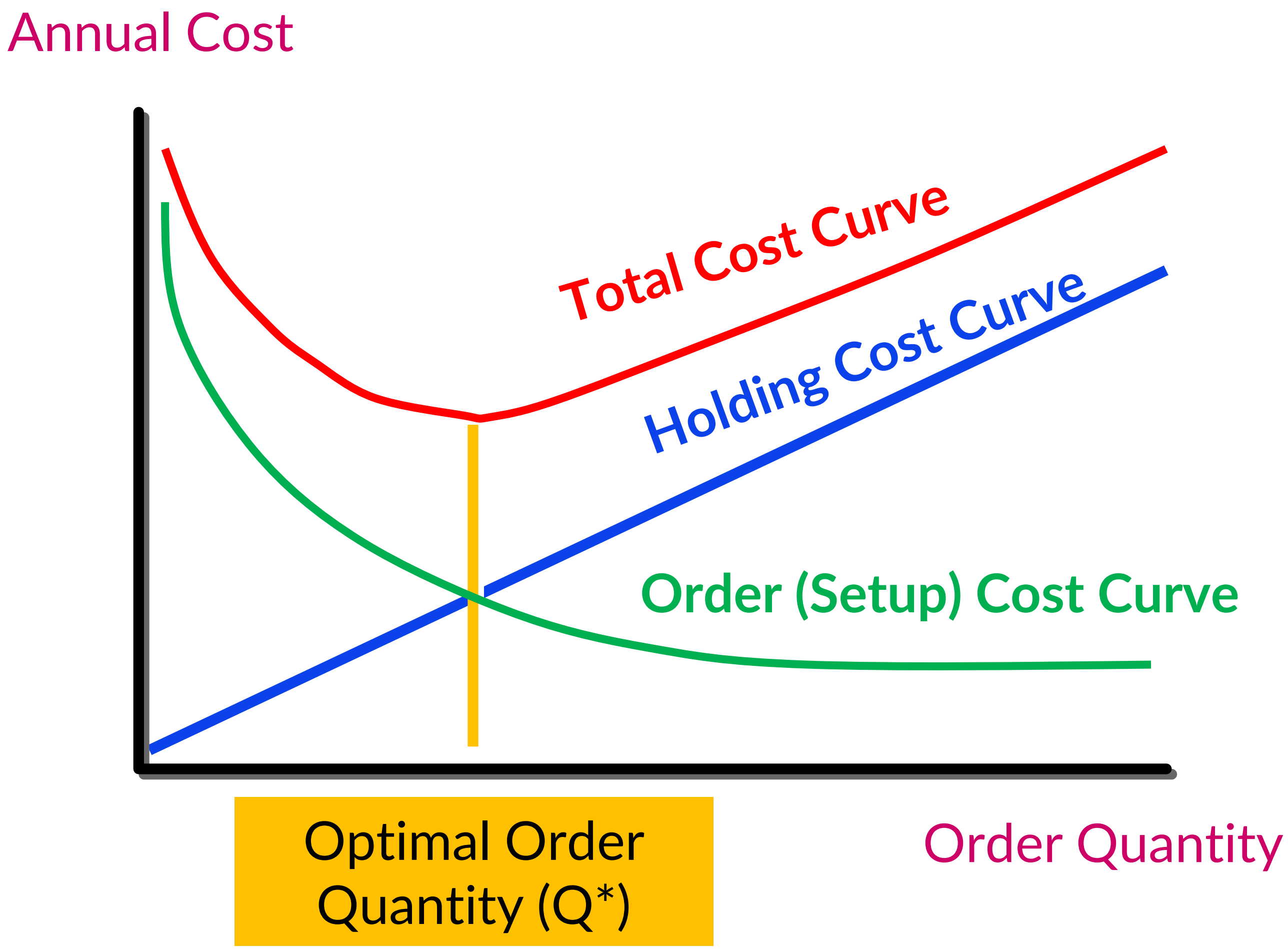
- Given we order Q each time, the per unit time cost is

$$\text{Total cost} = H \times Q/2 + S \times D/Q$$

- This holds for any given order quantity Q
- The optimal order quantity should minimize the total cost

- Trade-off: if we increase the ordered inventory Q each time
- We decrease the ordering cost ($S \times D/Q$) as we order fewer times
- But we increase the holding cost ($H \times Q/2$) as we maintain more inventory

EOQ Model: How Much to Order?



Q* balances the setup costs with inventory holding costs

Optimal Order Quantity

$$\text{Total cost} = TC(Q) = \frac{D \times S}{Q} + \frac{Q \times H}{2}$$

$$\frac{d\,TC(Q)}{d\,Q} = -\frac{D \times S}{Q^2} + \frac{H}{2} = 0$$

$$Q^* = \sqrt{\frac{2 \times D \times S}{H}}$$

Economic Order Quantity

EOQ Model: Optimal Order Quantity

$$Q^* = \sqrt{\frac{2 \times D \times S}{H}}$$

- Optimal Order Quantity = $\sqrt{\frac{2 \times \text{Order Cost} \times \text{Demand Rate}}{\text{Holding Cost}}}$

- It increases in order cost S and demand rate D, decreases in the holding cost H

- Square-root relation: if order cost (or demand rate) increases by four folds, the optimal order quantity increases by two folds

EOQ Model: Optimal Total Cost

$$TC(Q^*) = \frac{D \times S}{Q^*} + \frac{Q^* \times H}{2} = \sqrt{2 \times D \times S \times H}$$

- Total cost = $\sqrt{2 \times \text{Order Cost} \times \text{Demand Rate} \times \text{Holding Cost}}$
- Total Cost includes both fixed order cost (first term) and inventory holding cost (second term); measured for per unit of time
- It increases in order cost S, demand rate D, and holding cost H (square root relation)

EOQ Model: Optimal Holding and Ordering Cost

$$\text{Order Cost}(Q^*) = \frac{D \times S}{Q^*} = \sqrt{\frac{D \times S \times H}{2}}$$

$$\text{Holding Cost}(Q^*) = \frac{Q^* \times H}{2} = \sqrt{\frac{D \times S \times H}{2}}$$

- Both are measured for per unit of time
- At optimal order quantity, the ordering cost and holding cost equal each other; each contributes a half of the total cost

EOQ Model: Optimal Cost Per Unit

$$\frac{TC(Q^*)}{D} = \sqrt{\frac{2 \times H \times S}{D}}$$

- Cost Per Unit = $\sqrt{\frac{2 \times \text{Order Cost} \times \text{Holding Cost}}{\text{Demand Rate}}}$
- Cost per unit increases in order cost S and holding cost H (square relation)
- Cost per unit decreases in total demand rate: economy of scale in EOQ model
- Intuition: more demand reduces the average order cost

EOQ Model: Inventory Turnover

$$\frac{D}{Q/2} = \sqrt{\frac{2 \times D \times H}{S}}$$

- Inventory turn = $\sqrt{\frac{2 \times \text{Demand Rate} \times \text{Holding Cost}}{\text{Order Cost}}}$
- Inventory turnover increases in demand rate, holding cost, but decreases in order cost
- Intuition: if holding cost is high or ordering cost is low, we should "turn" the inventory more frequently

The M&S Example

- M&S has a stable demand for a line of computer it offers. Each week there is a demand for 100 computers. M&S incurs a fixed cost of $5000 every time it places an order. The marginal cost of a computer is $400, and the shop's holding cost of inventory is approximately 1% per week

$D = 100$, $S = 5000$, $H = 0.01 \times 400 = 4$ (convert to $ amount)

Optimal order size $Q = Q_{EOQ} = \sqrt{2 \times 5000 \times 100/4} = 500$ Units

Time between orders = Q/D = 500/100 = 5 weeks

Weekly ordering cost = (D/Q)S = (100/500) × 5000 = $1000

Weekly holding cost = (Q/2)H = (500/2) × 4 = 1000

Average flow time T = (Q/D)/2 = 500/100/2 = 2.5 weeks

Adding Per-unit Purchasing Cost

Sometimes we also consider the purchasing cost

Given:

D - Demand per unit of time (year, week, …)

S - Setup or Order Cost ($/setup; $/order)

H - Marginal holding cost ($/per unit per unit of time)

C - Purchasing cost ($/unit)

With purchasing cost, we have the following:

- Holding cost = $H \times Q/2$ per unit time
- Ordering cost = $S \times D/Q$ per unit time
- Purchasing cost = $D \times C$ per unit time
- Total cost = $DC + HQ/2 + SD/Q$ per unit time

Adding Per-unit Purchasing Cost

With Total cost = DC+HQ/2 + SD/Q per unit time, we note that

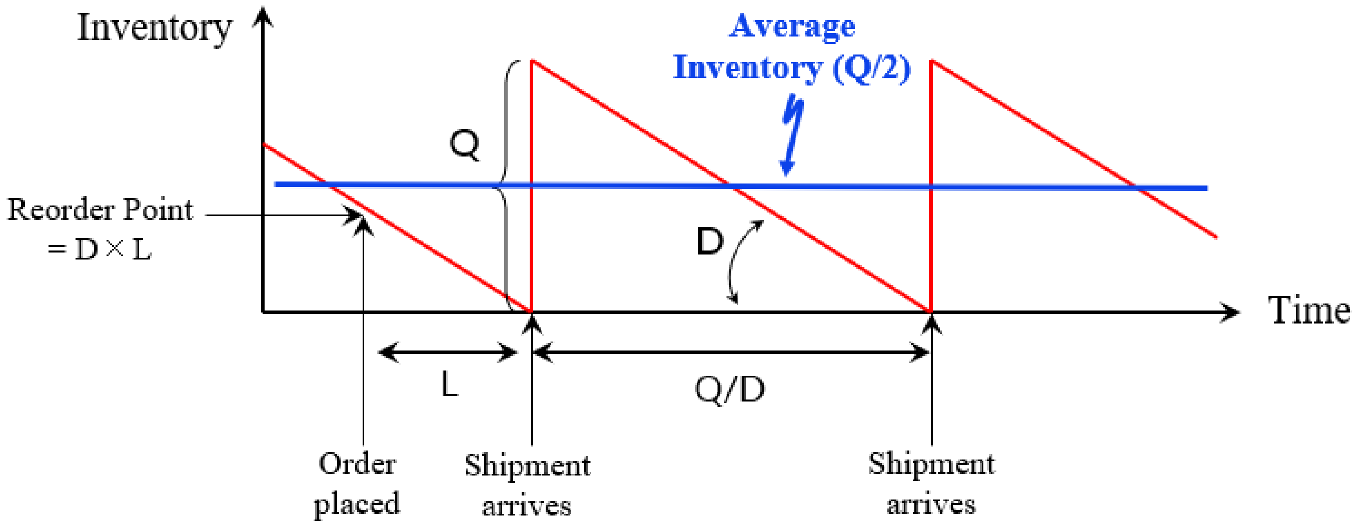
- The purchasing cost DC is a constant that does not depend on the order decision Q
- This is because the demand rate is fixed
- Thus, the optimal ordering plan is not affected

$$Q^* = \sqrt{\frac{2 \times D \times S}{H}}$$

- When the question asks you to compute "the total cost with purchasing cost included", you need to include the term DC

Lead Time in EOQ model

- Recall lead time is the time between placing the order and receiving the order



- The lead time does not affect our inventory pattern, thus does not change the total cost expression
- EOQ formula determines how much to order
- Lead time determines when to order

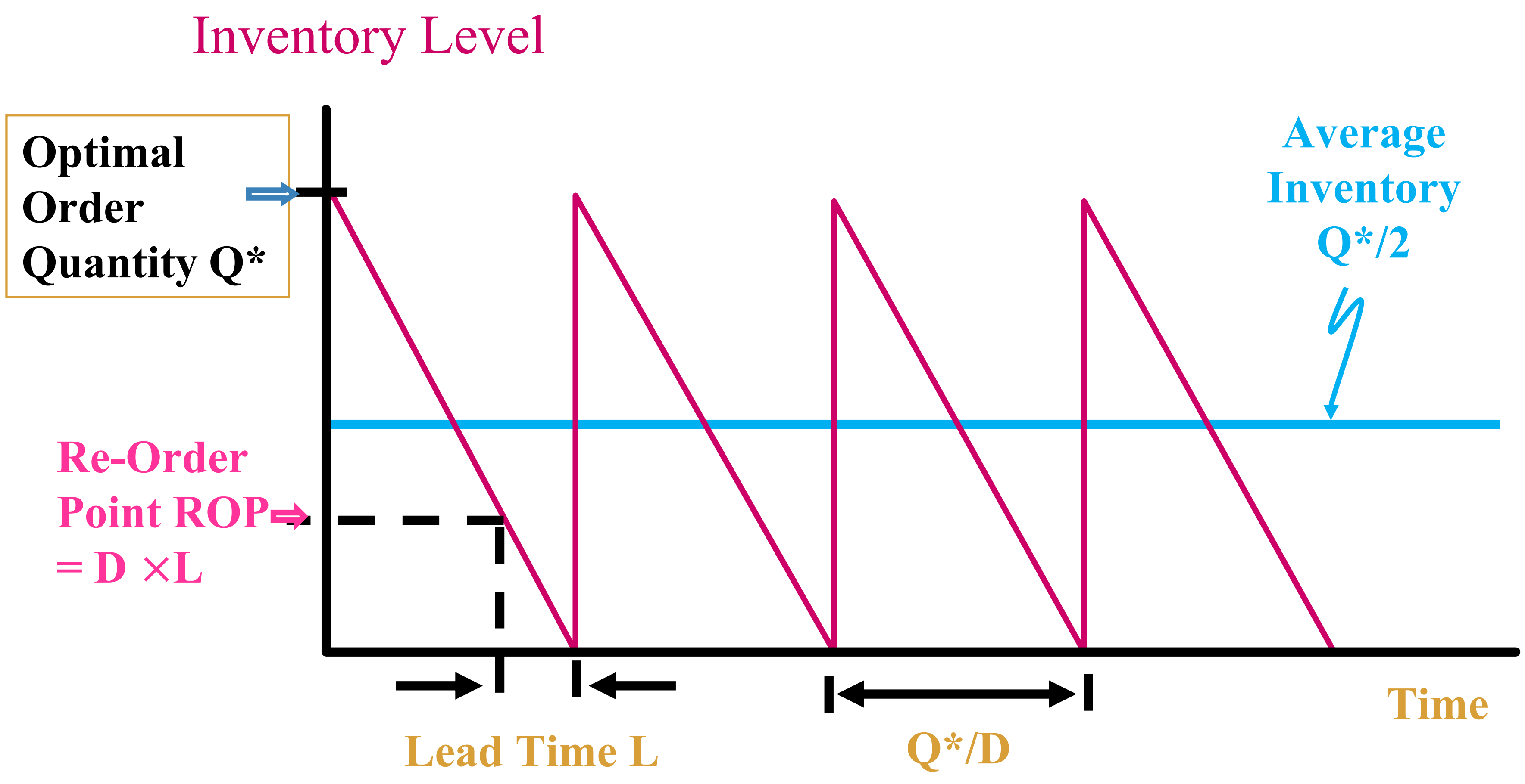
Lead Time in EOQ model

- Re-Order Point (ROP): the remaining inventory level such that we place a new order
- ROP = D×L: demand during the lead time
  - This holds when ROP < Q^* (always assumed in this course)

(M&S Store revisited) It takes one week from the time M&S places an order to the time the order is received. When should the store place an order?

- Given demand rate D = 100 and lead time L = 1, the ROP = 100 ×1 = 100 units
- M&S should place an order when the on-hand inventory drops to 100

EOQ Model: Summary



EOQ Model Equations

- Total cost per year: $TC(Q) = D \times C + \frac{D \times S}{Q} + \frac{Q \times H}{2}$
- Optimal order quantity (or EOQ)

$$Q^* = \sqrt{(2 \times D \times S)/H}$$

- Number of orders per year $N = D/Q^*$
- Reordering point: $ROP = D \times L$
  - This holds when $ROP < Q^*$ (assumed in this course)
- Order Q^* when inventory drops to ROP; then new orders are received when the inventory decreases to zero
- L affects the ordering time, but not ordering quantity and frequency; C does not affect our solution

D = Demand per year
S = Setup cost per order
H = Holding cost
C = Purchasing cost
L = Lead-time

EOQ Example

George Heinrich uses 18,000 units per year of a certain component. The per unit cost of this component is $500. Each order costs George $100. He operates 360 days per year and has found that an order must be placed with his supplier 4 days before he can expect to receive that order. The holding cost is $5 per unit per year. Find

- Economic order quantity

$$Q^* = \sqrt{\frac{2 \times 18000 \times 100}{5}} = 848.5\ units$$

- Annual holding cost and annual ordering cost

$$\frac{Q^* \times H}{2} = \$2121.3, \qquad \frac{D \times S}{Q^*} = \$2121.3$$

- Reordering point

$$ROP = \frac{18000}{360} \times 4 = 200\ units$$

EOQ Example

Suppose during the second year, demand has doubled, that is, George needs to use 36,000 units per year. Find

- Economic order quantity

$$Q^* = \sqrt{\frac{2 \times 36000 \times 100}{5}} = 1200\ units$$

- Annual holding cost and annual ordering cost

$$\frac{Q^* \times H}{2} = \$3000, \qquad \frac{D \times S}{Q^*} = \$3000$$

Knowledge Points

- Assumptions for EOQ model
- Definitions of demand rate, order cost, holding cost, purchasing cost, lead-time, etc
- Calculation of optimal order quantity, order frequency, reorder point
- Calculation of costs (holding, ordering, purchasing) at optimal order quantity or other order quantities

Takeaways

- An EOQ model is applicable when:
  - Demand is stable over time
  - Inventory is long-life, non-perishable (no obsolescence and spoilage)
  - Examples: groceries, chemicals, heavy industrial equipment, etc.
  - Fixed order cost and proportional inventory cost
- Trade-off: set-up cost and holding cost
- Solution: Optimal order quantity and frequency, as well as reordering point with lead time

ISOM 2700: Operations Management

Session 16. Inventory Management II: Newsvendor Model

Yiwen Shen
Dept. of ISOM, HKUST
Fall 2025, HKUST

Newsvendor Problem

- Every morning, a newsstand purchases newspapers to sell in the day

- Key feature: newspapers become outdated by the end of day!

- Trade-off: How many newspapers to purchase?
  - If too many, loss from unsold items
  - If too few newspapers, loss from potential sales

Structure of Newsvendor Problem

- Demand is uncertain (random)

- One-time ordering opportunity for supply (inventory quantity) before observing the demand

- Perishable goods: leftovers are costly
  – excess inventory will have lower (even zero) value

Newsvendor Model is Everywhere

- Christmas trees and decorations
  - Large depreciation after the holiday season
- ISOM 2700 classes
  - Financial loss if the classes are not fulfilled
- Operating room capacity for surgery
  - Reserving too much OR time would be wasteful

- Uncertain demand
- Orders before demand is realized (pre-commitment)
- Perishable product/capacity

Another Example: New Year Pocket

EOQ Model vs Newsvendor Model

	EOQ Model	Newsvendor Model
Product lifecycle	Long	Short
Number of periods	Multiple	Single
Demand	Deterministic	Random
Order opportunity	Multiple	Once

Key Decision: Inventory Level

- Key decision to make: how many inventory should we order in the newsvendor setting?
  - How many Christmas trees to buy before the holiday?
  - How many ISOM 2700 sessions to open in a semester?
  - How long to reserve in the OR for a surgery?

- This type of decision is very important in reality and has broad applications!

- However, it is not easy as we face random demand + costly leftovers

Probability Distribution of Random Demand

- The random demand is described by a probability distribution (discrete or continuous)

		Probability density	Cumulative probability
Demand	Frequency (out of 2000)	Prob(Demand = n)	Pr(Demand <= n)
10	89	4.45%	4.45%
20	183	9.15%	13.60%
30	260	13.00%	26.60%
40	387	19.35%	45.95%
50	467	23.35%	69.30%
60	289	14.45%	83.75%
70	180	9.00%	92.75%
80	92	4.60%	97.35%
90	32	1.60%	98.95%
100	21	1.05%	100.00%

- Make sure you distinguish between the probability density and cumulative probability

Service Level

- Throughout this session, we use D to denote the random demand, Q or n for the inventory level; Pr for the probability

- Service level: the probability that the store can fulfill the demand with its inventory

$$\text{Service Level} = \Pr(D \leq Q)$$

- This is an important operational target: managers usually want to maintain a higher service level
    - i.e., every customer who comes is able to get one

Inventory and Service Level

- Service level: the probability that the store can fulfill the demand with its inventory

$$\text{Service Level} = \Pr(D \le Q)$$

- The service level monotonically increases in the inventory level Q from 0% to 100%

- That is, if we want a higher service level, we should order more inventory

Probability of Shortage

- Shortage: some customers who want to buy cannot get the product as we do not have enough inventory

- It is given by the probability that demand is greater than the inventory

$$\text{Prob of Shortage} = 1\text{- Service Level} = \Pr(D > Q)$$

- Clearly, the probability of shortage decreases in the inventory level Q

Maximizing Expected Profit

- If we want higher service level, we should order more inventory
- However, this may hurt the expected profit, as we will have many leftover inventory at the end of period
- If we want to maximize our expected profit, what inventory level Q should we set?
- This is equivalent to setting a service level $\Pr(D \leq Q)$ given the demand distribution

Optimal Inventory Level

- We want to find the optimal inventory level that maximizes the expected profit from selling the product

- Trade-off in the inventory decision
    - ordering too few: unable to gain profit due to lost-sales
    - ordering too much: loss due to leftover inventory

- Question: what is the correct ordering quantity?

Marginal Analysis

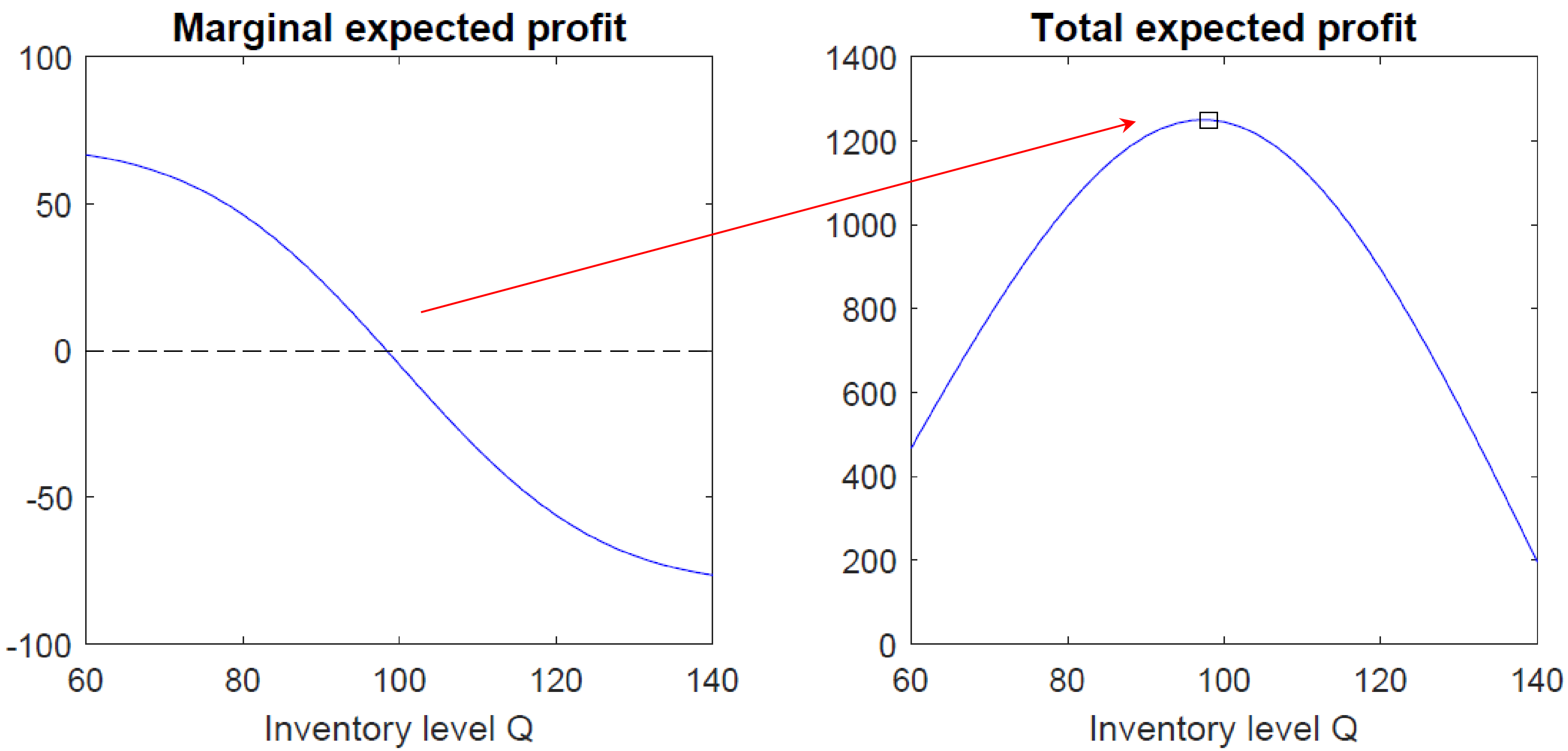
- We use marginal analysis to solve the optimal inventory level
  - A powerful tool in OM, Economics, and Finance

- Compute the marginal expected profit from the (Q+1)th unit inventory, given we already have ordered Q units

- If marginal profit is positive, then adding one more unit increases the total profit: you should increase it
- If marginal profit is negative, then adding one more unit decreases the total profit: you should decrease it

Marginal Profit and Total Profit

- The optimal solution is "usually" obtained when the marginal profit equals to zero
  - do not want to move in either direction



Newsvendor Example 1: Pumpkin

- You have the following demand forecast for Halloween:
  - $D \leq 200$ with probability 0.5
  - $D \leq 300$ with probability 0.9
- Price and costs:
  - The retailer buys pumpkins from a wholesaler at unit cost \$2
  - The retailer sells pumpkins at unit price \$5 (profit of 5-2=\$3)
  - Any unsold pumpkins has salvage value \$1 (loss of 2-1=\$1)
- Question: should we order one more unit, given we have ordered 200/300 pumpkins?

Marginal Expected Profit: Example

- Suppose the retailer has bought 200 pumpkins, we check the marginal expected profit of the 201st one

- This depends on the following factors:
  - If the demand is large enough, the 201^{st} pumpkin can be sold, then you earn \$3 (=5-2) dollars
  - If the demand is not large enough, the 201^{st} pumpkin cannot be sold, then you lose \$1 (=2-1) dollar

- Key question: what is the probability that the demand is large enough?

Marginal Expected Profit: Example

- If the retailer buys the 201st pumpkin, she needs to pay unit cost = \$2
- Two demand scenarios to be considered:
  (1) $D \leq 200$ (prob. 0.5): unable to sell the 201st pumpkin
  - Utility =\$1 – \$2= – \$1

  (2) $D > 200$ (prob. 0.5): able to sell the 201st pumpkin
  - Utility =\$5 – \$2= \$3
- That is, lose \$1 with prob 50%, earn \$3 with prob. 50%
- Expected value of the 201st pumpkin
  $= 0.5 \times (-\$1) + 0.5 \times \$3 = \$1 > 0$
- The retailer should buy the 201st pumpkin

Marginal Expected Profit: Example

- Suppose the retailer has bought 300 pumpkins

- Should the retailer buy the 301^{st} pumpkin?
  - She needs to pay unit cost = \$2

  (1) $D \leq 300$ (prob. 0.9): unable to sell
    - Utility =\$1 – \$2= – \$1

  (2) $D > 300$ (prob. 0.1): able to sell
    - Utility =\$5 – \$2= \$3
  - Expected value of the 251st pumpkin
    $= 0.9 \times (-\$1) + 0.1 \times \$3 = -0.6 < 0$

- Should the retailer buy the 301st pumpkin?
  - No, because the expected value is negative!

Newsvendor: Intuition

- As we have more inventory, the probability of selling another additional unit becomes smaller
  - Because it requires a larger demand level

- Thus, the expected value of the additional unit becomes smaller
  - More likely to suffer the loss due to leftover

- This means we should order up to a certain level such that the expected value drops (close) to zero

Marginal Analysis: General Case

- Suppose you have decided to stock/produce Q units
- Consider the incremental decision to stock/produce one more, namely a Q+1st unit

- Expected Profit of $(Q \rightarrow Q+1)$

= Profit if sold × Pr(it is sold) – Loss if unsold × Pr(it is unsold)

= (retail price – purchasing cost) × $\Pr(D > Q)$

– (purchase cost – salvage value) × $\Pr(D \leq Q)$

Critical Fractile Analysis

- We introduce two fundamental concepts in newsvendor models
- Understocking cost: C_u
  - Penalty of not having the unit when demand is sufficient
  - Marginal profit when n+1st unit is sold
  - Inventory example: Cu = retail price – purchase cost
- Overstocking cost: C_o
  - Penalty of having the unit when demand is insufficient
  - Marginal loss when n+1st unit is not sold
  - Inventory example: Co = purchase cost – salvage value

Understocking and Overstocking Costs

- Understocking cost: how much you lose for a unit when demand is greater than inventory
  - What you could have earned from lost-sales
  - e.g., I could have sold this product and earned $10 from it
- Overstocking cost: how much you lose for a unit when demand is smaller than inventory
  - Leftovers are costly because they are worth less
  - e.g., I lose $5 from each Christmas tree that I cannot sell
- In different questions, you need to find the correct Cu and Co based on the specific set-up (sometimes tricky)

Understocking and Overstocking Costs

- We can express the marginal expected profit of (Q+1)st unit as:

- Expected Profit of (Q → Q+1)

= (retail price – purchasing cost) × Pr(D > Q)

– (purchase cost – salvage value) × Pr(D ≤ Q)

$= C_u \times [1 - Pr(Demand \leq Q)] - C_o \times Pr(Demand \leq Q)$

$= C_u - (C_u + C_o) \times Pr(Demand \leq Q)$

- The marginal expected profit decreases with Q from positive (+Cu) to negative ($-C_o$)

Critical Fractile Analysis: Solution

- Remember the optimal inventory level should satisfy: Profit($Q \rightarrow Q+1$) = 0
- So, Profit of ($Q \rightarrow Q+1$) = $C_u - (C_u + C_o) \times Pr(Demand \le Q)$ equals zero when

$$Pr(Demand \le Q) = \frac{C_u}{C_u + C_o}$$

- The right-hand side is called: critical fractile
- Interpretation: Find Q such that the probability of demand smaller than it equals to the critical fractile
  - The optimal quantity Q^* that maximizes expected profits

Properties of the Optimal Solution

$$\Pr(\text{Demand} \leq Q^*) = \frac{C_u}{C_u + C_o}$$

Situation	$C_u \uparrow$	$C_o \uparrow$
Inventory Level	Increase	Decrease

- Higher understocking cost encourages manager to order more inventory
- … as he is penalized more for not having enough inventory

- Higher overstocking cost makes manager to order less inventory
- …as he suffers larger loss if he cannot sell the inventory

Expected Profit and Service Level: Trade-off

- In general, managers prefer both higher service level and higher expected profit

- Increasing inventory level always increases the service level, but it can decrease the expected profit due to leftover

- Thus, we face a trade-off between expected profit and service level, which is controlled by the actual order quantity

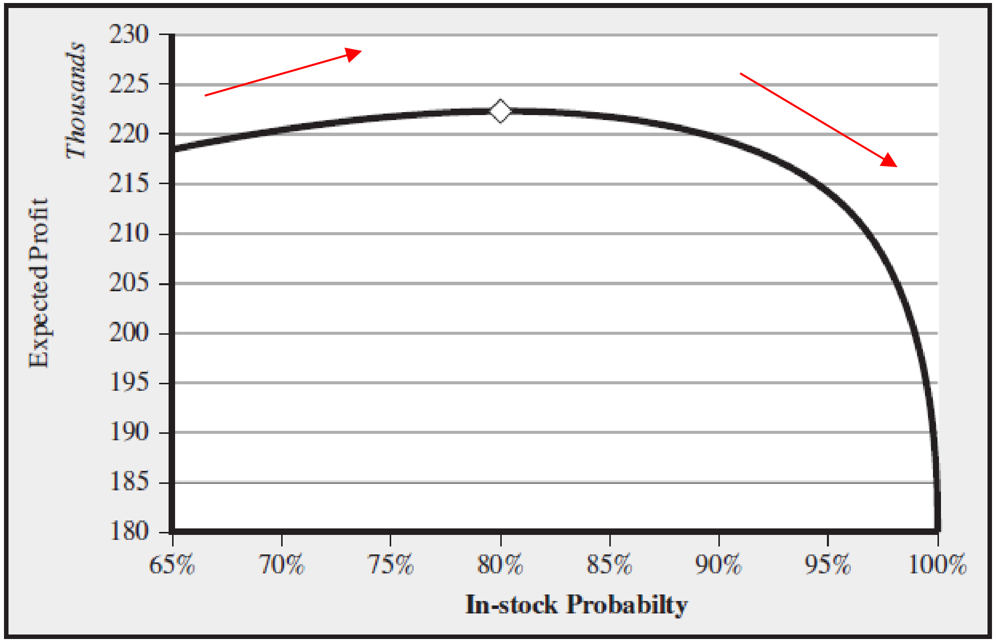
- In practice, we need to find a balancing point between profit and service level

Expected Profit and Service Level

- When current inventory is lower than the optimal level from the newsvendor solution, if we order more:
  - increase expected profit as its expected margin profit is positive
  - increase the service level as we have more inventory

- When current inventory is higher than the optimal level from the newsvendor solution, if we order more:
  - decrease expected profit as its expected margin profit is negative
  - increase the service level as we have more inventory
  - A trade-off appears!

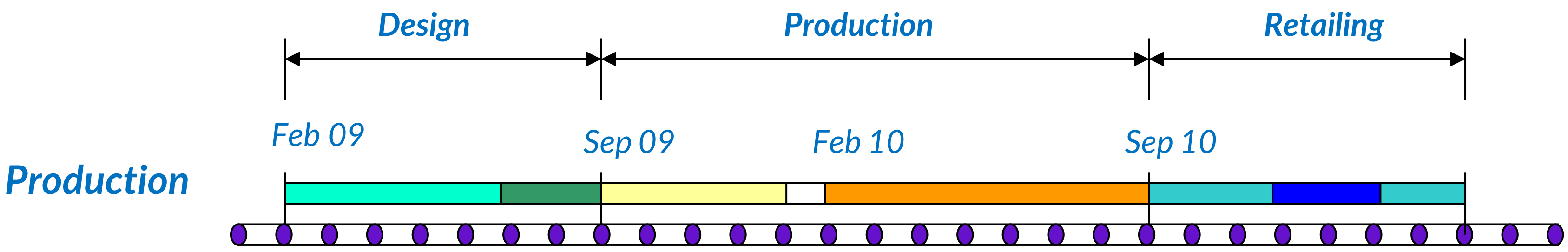
Plot: Expected Profit Against Service Level



An example from Cachon and KC's book

- The solution from the newsvendor model corresponds to the highest point (largest expected profit)
- As we further increase the inventory level, we have a higher service level but lower expected profit

Discrete Demand: SnowTime Sports



- Properties of SnowTime Sporting Goods:
  - Prices drop after winter season (perishable)
  - One production/ordering opportunity
  - A probabilistic forecast based on past sales, industry knowledge, and economic conditions

SnowTime Revenue and Cost Data

- For each unit
  - c = production cost = \$80
  - r = selling price = \$125
  - s = salvage value of unsold unit = \$20
- Recall
  - C_u = marginal profit of extra unit sold
  - C_o = marginal cost of unsold unit
- Hence, understocking and overstocking costs are:
  - $C_u = r – c = \$45$ and $C_o = c – s = \$60$

SnowTime Demand: Discrete Distribution

Demand	Probability
8,000	0.11
10,000	0.11
12,000	0.28
14,000	0.22
16,000	0.18
18,000	0.10

- Suppose the store faces discrete demand levels with the above distribution
- The table gives the probability that the demand equals a particular level $\Pr(D = Q)$
- In this case, we only consider the discrete demand levels as potential inventory choices

Cumulative Demand Probability

Demand	Probability
8,000	0.11
10,000	0.11
12,000	0.28
14,000	0.22
16,000	0.18
18,000	0.10

Demand	Probability	Cumulative probability
8,000	0.11	0.11
10,000	0.11	0.22
12,000	0.28	0.50
14,000	0.22	0.72
16,000	0.18	0.90
18,000	0.10	1.00

- We need to first calculate the cumulative probability: the probability that the demand is equal to or smaller than a certain level
- Then, we can use the critical fractile in newsvendor model

Optimal Solution for SnowTime

- The optimal production quantity Q^* satisfies

$$\Pr(\text{Demand} \leq Q^*) = \frac{C_u}{C_u + C_o} \approx 0.429$$

Demand	Probability	Cumulative probability
8,000	0.11	0.11
10,000	0.11	0.22
12,000	0.28	0.50
14,000	0.22	0.72
16,000	0.18	0.90
18,000	0.10	1.00

It falls between 10,000 and 12,000.
By round-up rule, we take $Q^* = 12{,}000$
This is sufficient for this course.

The most rigorous way is to compare the expected profit for the levels below and above the target, but we do NOT need this for the course. We simply take the upper side.

EOQ Model vs Newsvendor Model

	EOQ Model	Newsvendor Model
Product lifecycle	Long	Short
Number of periods	Multiple	Single
Demand	Deterministic	Random
Order opportunity	Multiple	Once

Ojective	Minimize total cost	Maximize profit
Cost	Holding cost and order cost	Overstocking cost and understocking cost
Method	First order derivative of cost function	Marginal analysis of expected value

Knowledge Points

- Assumptions and set-up for newsvendor model
- Definitions of over-stocking, under-stocking cost, and critical fractile
- Idea of marginal analysis for determining optimal order quantity
- Solving optimal order quantity and service level for discretized demand distribution and normal demand distribution

ISOM 2700: Operations Management

Session 17. Inventory Management III: More on Newsvendor Model

Yiwen Shen
Dept. of ISOM, HKUST
Fall 2025, HKUST

Agenda

- **Newsvendor model: Ordering level under normal demand**

- Newsvendor model: performance metrics under normal demand

- Benefit of risk-pooling: evidence from newsvendor model

Recap: Discrete Demand Distribution

- We have learnt that in the newsvendor model, the profit-maximizing inventory level should satisfy:

$$\text{Prob}\{D \leq Q\} = C_u/(C_u + C_o)$$

- We have also learnt how to get the optimal Q when the demand follows a discrete distribution

- Remember to first get the cumulative distribution function Prob{D ≤ Q}
- Then, find the interval that contains the critical fractile

Normally Distributed Demand

- We now consider the case where customer demand D is a normal random variable

- The demand has mean μ and standard deviation σ

- The normal demand follows a continuous distribution
  - We ignore the possibility of negative demand

- Question: how do we solve the corresponding quantities under normal demand?

Find Service Level Given Inventory

- Suppose demand D is normally distributed with mean of μ and standard deviation of σ
- The current inventory is Q, what is the service level, i.e., Prob(D ≤ Q)?

- Remember for normal random variable D, we have for any Q:

$$\text{Prob}\{D \leq Q\} = \text{Prob}\left\{\frac{D-\mu}{\sigma} \leq \frac{Q-\mu}{\sigma}\right\} = \text{Prob}\left\{Z \leq \frac{Q-\mu}{\sigma}\right\}$$

where Z is a standard normal random variable with mean zero and standard deviation one (check Session 4)

- We can then find the probability $\text{Prob}\left\{Z \leq \frac{Q-\mu}{\sigma}\right\}$ using the standard distribution table

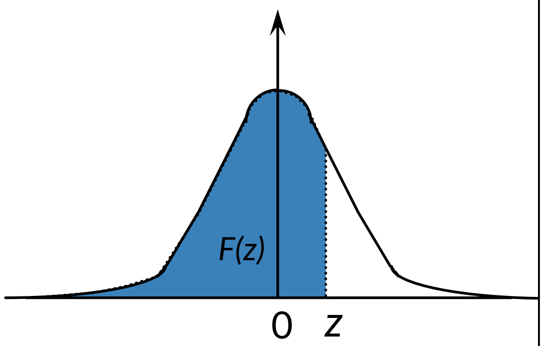
Find Inventory Given Service Level

- Suppose demand D is normally distributed with mean of μ and standard deviation of σ
- If we want to achieve a service level of α, what inventory level should we set?
- We want $\text{Prob}\{D \leq Q\} = \text{Prob}\left\{\frac{D-\mu}{\sigma} \leq \frac{Q-\mu}{\sigma}\right\} = \alpha$
- We first find z such that $\text{Prob}\{Z \leq z\} = \alpha$ using the standard distribution table. Then we set $Q = \mu + \sigma \times z$
- This is the desired inventory level as:

$$\text{Prob}\{D \leq Q\} = \text{Prob}\left\{\frac{D-\mu}{\sigma} \leq \frac{Q-\mu}{\sigma}\right\} = \text{Prob}\{Z \leq z\} = \alpha$$

The Standard Normal Distribution

$F(z) = \text{Prob}(Z \leq z)$



$F(z)$ = 98.7% implies $z^* = 2.23$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997

Reading Standard Normal Distribution Function Table

- Q: What is the probability the outcome of a standard normal will be z = 0.28 or smaller?
  - Look for the intersection of the fourth row (with the header 0.2) and the ninth column (with the header 0.08) because 0.2+0.08 = 0.28, which is the z we are looking for.
- A: The answer is 0.6103, which can also be written as 61.03%

Standard Normal Distribution Function Table (continued), Φ(*z*)

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224

Reading Standard Normal Distribution Function Table

- Q: For what z is there a 60% chance that the outcome of a standard normal will be that z or smaller?
  - The value of z falls between 0.25 and 0.26

Standard Normal Distribution Function Table (continued), $\Phi(z)$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224

- Round-up rule: we take the larger value of z, i.e., 0.26 here
  - Similarly, we choose the larger order quantity if the solution falls between two integers

Normally Distributed Demand: Examples

Suppose the customer demand follows a normal distribution with mean of 500 and standard deviation of 60.

Q1: If the retailer orders 580 products. What is the service level?

The z-score is (580-500)/60 = 1.333, using the distribution table, the service level is $Pr(Z \leq 1.33)$= 90.82%

Q2: If the retailer wants a service level of 90%, how many products should she order?

The corresponding z-score is z = 1.29 (round-up), thus the inventory is Q= 500+60*1.29= 578 units

Optimal Q^* with Normal Demand

- We now solve the optimal ordering quantity Q when customer demand D is a normal random variable
  - With mean μ and standard deviation σ
  - Ignore the possibility of negative demand

- Again, we want to find the Q^* such that

$$\text{Prob}\{D \leq Q^*\} = C_u/(C_u + C_o)$$

- Note that the C_u and C_o do not depend on the demand distribution
  - They are determined by the costs and profits

Optimal Q^* with Normal Demand

- That is, we want to find Q^* to achieve the service level $C_u/(C_u + C_o)$

- First, use the distribution table to find the optimal z-score z^* that satisfies

$$\text{Prob}\{Z \le z^*\} = C_u/(C_u + C_o)$$

- Then, we solve the optimal $Q^* = \mu + z^* \times \sigma$ with μ is the mean of distribution, σ is standard deviation

An Apparel Retailer

- The retailer can order a fashion T-shirt at the price of $60 per unit. The T-shirt can be sold at the price of $120 per unit, and has a salvage value of $20 per unit. The demand has a normal distribution with a mean of 100 units and standard deviation of 40 units (ignore the possibility of negative demand)

- C_u = 120 – 60 = $60 and C_o = 60 – 20 = $40
- Critical fractile = $C_u/(C_u+C_o)$ = 60/100 = 0.6

- Z-score from the $\Phi(z)$ Table: z = 0.26 (round-up rule)
- Optimal quantity = 0.26×40+100 = 110.4 ≈ 111 units

Steps for Solving NV Model

- Identify the overstocking and understocking cost:
  - Understocking cost: C_u = profit when a unit is sold
  - Overstocking cost : C_o = loss when a unit cannot be sold

- Calculate the critical fractile: $C_u/(C_u+C_o)$
  - This is also the service level Prob($D \leq Q$) with optimal inventory Q

- These two steps do not depend on the demand distribution

Steps for Solving NV Model

- Optimal order quantity: Discrete demand
  - Calculate cumulative probability for each demand level
  - Find the smallest inventory level with cumulative probability equal to/higher than critical fractile (round-up rule)

- Optimal order quantity: Normal demand
  - Find the corresponding z-score from distribution table
  - Set order quantity $Q^* = \mu + z^* \times \sigma$

Agenda

- Newsvendor model: Ordering level under normal demand
- **Newsvendor model: performance metrics under normal demand**
- Benefit of risk-pooling: evidence from newsvendor model

Performance Metrics in NV Model

- Let Q denote the inventory and D denote the realized demand
- Leftover = max{Q-D, 0}: number of unsold inventory due to insufficient demand
- Sales = min{Q, D}: number of sold items
- Lost-sales = max{D-Q, 0}: number of sales that you lost due to insufficient inventory
- With random demands, these quantities are also random

Performance Metrics: Examples

- Suppose you have ordered 120 units of inventory, and the realized demand is 130. Then,
  - Your leftover is zero: max(120-130,0) = 0
  - Your sales is 120: min(120,130)=120
  - Your lost sales is ten: max(130-120,0) = 10

- Suppose you have ordered 120 units of inventory, and the realized demand is 90. Then,
  - Your leftover is 30: max(120-90,0) = 30
  - Your sales is 90: min(120,90)=90
  - Your lost sales is zero: max(90-120,0) = 0

Realized and Expected Performance Metrics

- For a given realized demand level, only one of the leftover and lost-sales can be positive, the other must be zero
- However, both the expected leftover and expected lost-sales are positive, as we take expectation over all possible demand levels
- Therefore, make sure you distinguish between realized and expected metrics
- For example, if you toss a coin, only one side (head or tail) will appear in a trial
- But both sides have a 50% chance in expectation

Expected Leftover Inventory

- If demand is greater than Q, all units will be sold
- Otherwise, we will have some leftover inventory

- We compute the expected leftover inventory when we set the inventory at Q

- Under normally distributed demand, this can be computed as

$$\text{Expected leftover inventory} = \sigma \times I(z)$$

- Here $z = \frac{Q-\mu}{\sigma}$ is the z-score associated with Q; I(z) is the inventory function for standard normal variable
  - available from Standard Normal Inventory Function Table

Expected Leftover Inventory: Example

- In the apparel retailer example, suppose we use the optimal inventory level Q^*
- At optimal inventory, the z-score corresponding to critical fractile ratio is z = 0.26
- From the I(z) Table, we have: I(0.26) = 0.5424
- Then, the expected leftover inventory is:
  - $\text{Expected leftover} = \ 0.5424 \times 40 \ = 22 \text{ units}$
- Here we just round to nearest integer (this does not matter in exams, you always choose closest option)

Distribution Table vs Inventory Table

- The distribution table Φ(z) is for calculating Pr(N(0,1)≤z)
  - All values in distribution table are between 0 and 1
- The inventory table I(z) is for calculating expected inventory level
  - The values in the inventory table can be larger than 1

Standard Normal Distribution Function Table, Φ(z)

z	−0.09	−0.08	−0.07	−0.06	−0.05
−4.0	0.0000	0.0000	0.0000	0.0000	0.0000
−3.9	0.0000	0.0000	0.0000	0.0000	0.0000
−3.8	0.0001	0.0001	0.0001	0.0001	0.0001
−3.7	0.0001	0.0001	0.0001	0.0001	0.0001
−3.6	0.0001	0.0001	0.0001	0.0001	0.0001
−3.5	0.0002	0.0002	0.0002	0.0002	0.0002
−3.4	0.0002	0.0003	0.0003	0.0003	0.0003

Standard Normal Inventory Function Table, *I*(*z*)

z	−0.09	−0.08	−0.07	−0.06	−0.05
−4.0	0.0000	0.0000	0.0000	0.0000	0.0000
−3.9	0.0000	0.0000	0.0000	0.0000	0.0000
−3.8	0.0000	0.0000	0.0000	0.0000	0.0000
−3.7	0.0000	0.0000	0.0000	0.0000	0.0000
−3.6	0.0000	0.0000	0.0000	0.0000	0.0000
−3.5	0.0000	0.0000	0.0000	0.0000	0.0000
−3.4	0.0001	0.0001	0.0001	0.0001	0.0001

Expected Sales

- Sales: number of units actually sold

- Let Q denote the optimal order quantity, we have

$$\text{Sales} + \ \text{Leftover inventory} = \text{Q}$$

- Taking expectations on both sides, we can compute it from expected leftover inventory:

$$\text{Expected sales} = \text{Q} \ - \text{Expected leftover inventory}$$

- In the apparel retailer example, the expected sales is:

$$\text{Expected sales} = 111 \ - 22 = 89 \text{ units}$$

Expected Lost Sales

- Lost sales: units not sold because running out of inventory
  - It is positive when demand exceeds inventory

- It satisfies:

$$\text{Sales} + \ \text{Lost sales} = \text{Demand}$$

- Taking expectations on both sides, we can compute it from expected sales

$$\text{Expected lost sales} = \text{Expected demand} \ - \text{Expected sales}$$

- In the apparel retailer example, the expected lost sales is:

$$\text{Expected lost sales} = 100 \ - 89 = 11 \text{ units}$$

Expected Profit

- The expected profit includes both the units sold and the units leftover

$$\text{Profit} = [(\text{Price} - \text{Purchase Cost}) \times \text{Sales}] - [(\text{Purchase Cost} - \text{Salvage value}) \times \text{Leftover}]$$

- Taking expectations on both sides:

$$\text{Expected profit} = [(\text{Price} - \text{Purchase Cost}) \times \text{Expected sales}] - [(\text{Purchase Cost} - \text{Salvage value}) \times \text{Expected leftover}]$$

- By definition, the expected profit is maximized at the optimal order quantity
- In the apparel retailer example, the expected profit is:

$$\text{Expected profit} = 60 \times 89 - 22 \times 40 = \$4460$$

Newsvendor Model Metrics: Summary

- Leftover inventory and sales sum up to the order quantity
- Sales and lost sales sum up to the actual demand
- The expected leftover inventory can be directly computed from the z-score and the I(z) table
- The expected profit includes the marginal profit from units sold and the marginal loss from the units leftover
- The optimal order quantity maximizes the expected profit

Agenda

- Newsvendor model: Ordering level under normal demand
- Newsvendor model: performance metrics under normal demand
- **Benefit of risk-pooling: evidence from newsvendor model**

Apparel Retailer Example

- Suppose we have two identical apparel retailers as described in previous example. Their demands are independent.

- The retailer can order a fashion T-shirt at the price of $60 per unit. The T-shirt can be sold at the price of $120 per unit, and has a salvage value of $20 per unit.

- The demand has a normal distribution with a mean of 100 units and standard deviation of 40 units.

- We've solved this problem by newsvendor model, and computed various performance metrics

Separated vs Pooled

- Consider two inventory management modes for the two retailers

- Separated: the two retailers order and keep their inventory separately based on their own demand

- Each of them will use the optimal order quantity from the previous newsvendor model

- Pooled: set up a pooled inventory for the two retailers; order and keep the common inventory based on the total demand

Location Pooling

- Location pooling: combining the inventory from multiple *individual territories* into a single location

- One of the risk-pooling strategies for reducing variability
  - Location pooling, virtual pooling, product pooling, lead time pooling, and capacity pooling

- Several cautions must be considered for location pooling
  - Set-up cost of the shared inventory location
  - Impact on the efficiency of the sales
  - Incentive conflict of individual retailers/sales

Apparel Retailer: Separated Mode

- We compute the performance metrics for the apparel retailer example under the separated and pooled modes
- Separated mode: each retailer solves its own newsvendor model
- Critical fractile = $C_u/(C_u+C_o)$ = 60/(60+40) = 0.6
- Normal demand: mean = 100, standard deviation = 40
- Corresponding Z-score: z = 0.26

	Order quantity	Expected leftover	Expected sales	Expected lost sales	Expected profit
Each retailer	111	22	89	11	4460
Total	222	44	178	22	8920

Pooled Mode: Demand

- We pool the demands and inventory for the two retailers, i.e., order one-time for the two retailers combined

- We first determine the distribution for the total demand
  - Sum of independent normal distribution is still normal
  - Mean: $100 \times 2 = 200$; Standard deviation = $40 \times \sqrt{2} = 56.57$

- The total demand follows a normal distribution with mean of 200 and standard deviation of 56.57

- Decrease in coefficient of variation (CV):

$$\text{CV}_{\text{sep}} = \frac{40}{100} = 0.40 \quad \text{and} \quad \text{CV}_{\text{pool}} = \frac{56.57}{200} = 0.28$$

Pooled Mode: Solution

- We solve the optimal order quantity under the pooled mode by newsvendor model
  - with the new demand distribution

- The underage and overage costs are unchanged (depending on the marginal profit/loss)
  - Critical fractile = $C_u/(C_u+C_o)$ = 60/(60+40) = 0.6
  - Z-score = 0.26 also does not change

- New optimal order quantity:

$$Q = 0.26 \times 56.57 + 200 = 215 \text{ units}$$

Pooled Mode: Performance Metrics

- We can compute the performance metrics under the pooled mode following the steps described before

- Expected leftover inventory = $0.5424 \times 56.57 = 31$ units
- Expected sales =$215 - 31 = 184$ units
- Expected lost sales = $200 - 184 = 15$ units

- Expected profit = $60 \times 184 - 40 \times 31 = \$9800$

Separated vs. Pooled

- We compare different performance metrics under the separated (sum of the two) and pooled modes:

	Order quantity	Expected leftover	Expected sales	Expected lost sales	Expected profit
Separated (total)	222	44	178	22	8920
Pooled	215	31	184	16	9800
Difference	-7	-13	+6	-6	+880

- Higher expected profit under location pooling

Benefits from Location Pooling

- With location pooling, we get the following benefits:
- The inventory level moves closer to the expected demand
- Fewer expected leftover inventory; fewer expected lost sales; more expected sales
- Higher expected profit
- These properties hold for general newsvendor models (their proofs are not required, but you might try)

Benefit of Pooling: Intuitions

- Reduced uncertainty in demand: we only need to care about the total demand, instead of each store's demand
- This can be seen from the smaller coefficient of variation

- Better match between supply and demand: this is similar to the common vs. separated queue comparison
- In the separated mode, it is possible that one store has leftover but the other store has lost-sales
- This is impossible in the pooled mode, as the inventory is shared among the two --- a hedging effect!

Benefit of Pooling: Other Factors

- (Not Required)
- There are many factors affecting the benefit of pooling. Some general observations are in below.

- The benefit of pooling is larger under higher demand variability

- The benefit of pooling is smaller when the two demands are more positively correlated
- Perfectly positively correlated ($\rho=1$): no benefit from pooling

- The marginal gain from pooling another store decreases in the number of stores already pooled (same as M/M/s queue)

Knowledge Points

- Performance metrics for newsvendor model at the optimal order quantity
  - Expected leftover inventory, sales, lost sales, and profit

- Tradeoff between profit and service level

- Benefit of pooling in newsvendor model:
  - Distribution of pooled demand
  - Model solution/performance metrics with and without pooling

ISOM 2700: Operations Management

Session 18. Revenue Management I

Yiwen Shen
Dept. of ISOM, HKUST
Fall 2025, HKUST

Agenda

- **Motivation to revenue management**

- Capacity-based revenue management

What is Revenue Management?

- Revenue Management (RM) was first developed in airline industry in 1980
- It is the practice of maximizing revenue by optimizing product availability (capacity) and price
- Evolution of RM

Capacity-based RM	Price-based RM
• Widely used in airlines, hotels, and car rental industries • Fixed and perishable resources • …	• Markdown pricing in retailing • Dynamic pricing in e-commerce • Surge pricing by Uber • …

Why Revenue Management?

Before RM

Revenues	**1000**
COGS	**800**
Gross Margin	**200**
Other Expenses	**150**
Net Margin	**50**

After RM

Revenues	**1020**
COGS	**800**
Gross Margin	**220**
Other Expenses	**150**
Net Margin	**70**

Change

Revenues	**+2%**
COGS	**0%**
Gross Margin	**10%**
Other Expenses	**0%**
Net Margin	**40%**

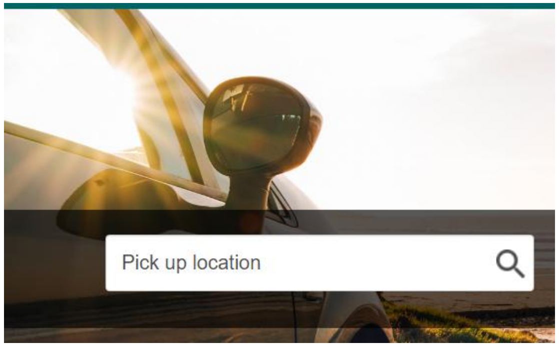
A small increase in revenue can largely increase profit

RM: Industrial Applications

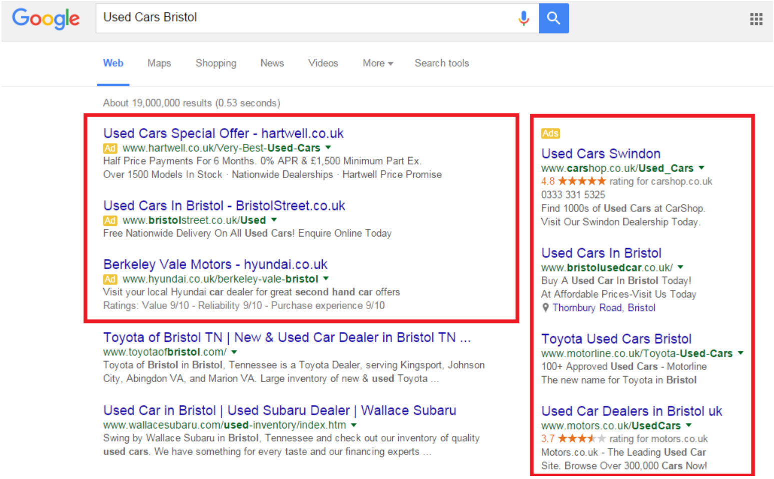
Airlines

Hotel

Car rental



Media & broadcasting



Retailing

Behind RM: Customer Segmentation

- Behind RM: customer segmentation and differentiation (e.g., price discrimination in economics)
  - time-based differentiation: e.g., regular selling season followed by markdown season
  - quality-based differentiation: e.g., different classes of airline tickets; regular ticket versus FastPass in Disney

- Goal: match supply with demand
  - sell right product to right customer at right time and price

- Intersection of demand, supply, and analytics
  - demand: data collection, forecasting
  - supply: capacity and service design
  - analytics: capacity control, dynamic pricing

Examples: Airline Classes

Sort By: # of Stops

	Departure	Arrival	Economy Super Saver	Economy Saver	Economy Flexible	Instant Upgrade	Business Flexible	First Special	First Flexible
AA Non-Stops									
AA 177	05:40 PM JFK	09:10 PM SFO	$479	$622	$954	$2491	$2590	Not Available	$3138
+ Flight Details									
AA with Stops									
AA 773	05:05 PM LGA	08:25 PM DFW	$479	$622	$954	$1749	Not Available	$2099	$2229
AA 1575	09:20 PM DFW	11:05 PM SFO							
+ Flight Details									

Agenda

- Motivation to revenue management
- **Capacity-based revenue management**

Capacity-based RM: Airline Industry

- A typical airline operates with 73% of its seats filled
  - but needs to fill 70% of its seats to break even
- Airline seats are highly perishable
  - once the plane takes off, empty seats will be wasted
- There are usually multiple classes of customers
  - Some are time sensitive; some are price sensitive
- Key decisions:
  - How many seats to make available for each customer class?

Application 1: Two-Class Allocation

- An airline has started to sell tickets for the flight from Hong Kong to London on May 29

- An airplane has 335 seats for passengers

- Two classes of passengers
  - Leisure: sensitive to price and buy tickets in advance
  - Business: price insensitive and buy at the last minute

Two-price Strategy

- To exploit the customer difference, the company usually uses a two-price strategy

- Two-price strategy:
  - Offer two prices, f_1 = \$7950 > f_2 = \$5250
  - Discount price (f_2) targets leisure travelers
  - Full price (f_1) targets business travelers

- Here we assume the prices are given
  - In reality, they are also optimized by the firm

Two-price Strategy: Assumptions

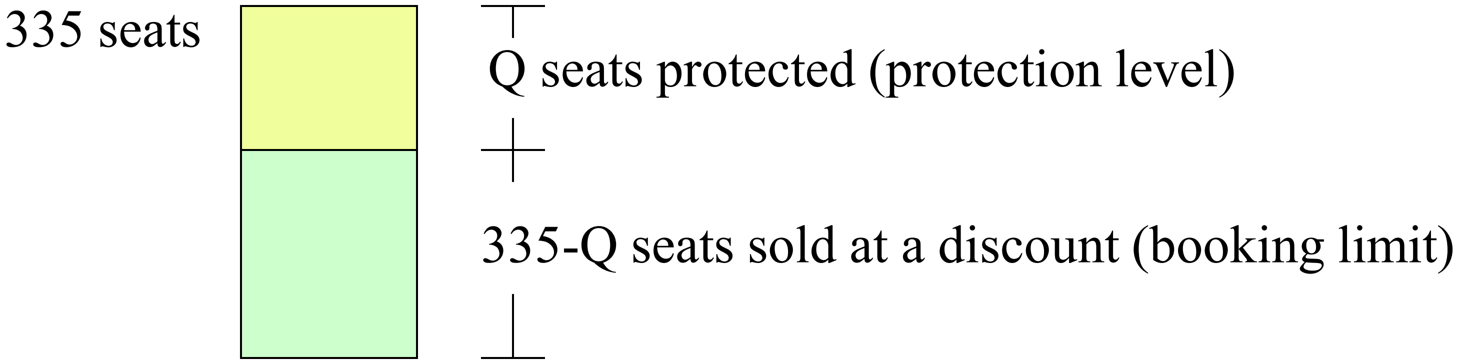
- Discount tickets are always sold out to leisure travelers in advance with sufficient demands
  - That is, you can sell as many discount tickets as you want

- Full price tickets are only sold to business travelers in the last minute, and there is only limited demand
  - Here we assume full tickets demand D_b follows a normal distribution with mean 25 and standard deviation of 5

- Trade-off: If you sell too many discount tickets in advance, you may lose the higher revenue from the full price tickets
- If you sell too few discount tickets, you may end up with empty seats

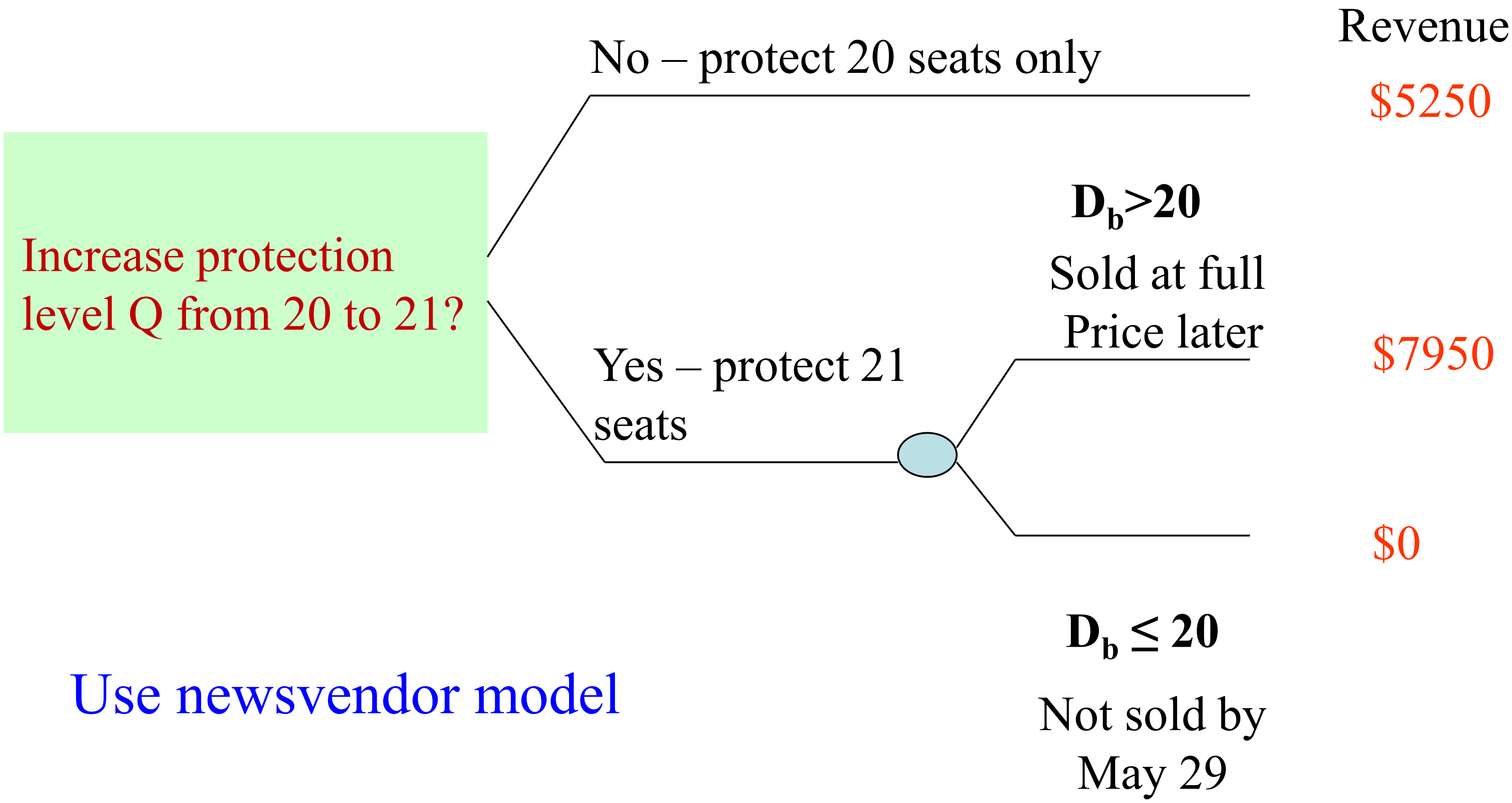
Two-Class Allocation Problem

- The goal is to maximize the expected revenue from the flight
- Decision variable: Protection level (Q)
  - How many seats (Q) should be reserved for business travelers?



- Booking limit: the maximum number of seats that may be sold at the discount price
  - Booking limit + protection level = total capacity

Marginal Analysis



Use newsvendor model

Convert to a Newsvendor Problem

- To use the newsvendor model, we first figure out what is "demand" and what is "inventory"
- In the two-class allocation problem, the "demand" is the number of business travellers
- The "inventory" is the number of seats reserved to them (protection level)

- Then, the overstocking cost is the penalty when you reserve one more seat but has no business traveller
- The understocking cost is the penalty when you reserve one fewer seat (sell it in advance) but has a business traveller eventually

Marginal Analysis

- Overstocking cost (too high protection level Q)
  - $C_o = f_2 = \$5250$ (loss when we cannot sell)
- Understocking cost (too little protection level Q)
  - $C_u = f_1 - f_2 = 7950 - 5250 = \$2700$ (benefit from selling at full price)
- Optimal protection (inventory) level satisfies

$$\Pr\left(D_b \leq Q\right) = \frac{C_u}{C_u + C_o} = \frac{2700}{2700+5250} = 0.339$$

- Here D_b is the number of business travelers
- Given D_b is Normal(25, 5), the protection level should be set as

$$Q^* = 25 + 5z^* \approx 23 \text{ seats}$$

with the z-score $z^* \approx -0.41$ from the distribution table

Two-class Problem: Performance Metrics

- In the two-class problem, the inventory level is the number of full tickets reserved
- The demand level is the number of business travelers

- Expected leftover: full-price tickets not able to sell at the end (empty seats)
  - It happens when inventory exceeds demand
  - Expected leftover = $I(z) \times \sigma = I(-0.41) \times 5 = 0.227 \times 5 = 1.135 \approx 1$

- Expected sales: full-price tickets that are sold to business travelers
  - Expected sales = Q – expected leftover = 23 – 1 = 22

Two-class Problem: Performance Metrics

- Expected lost-sales: business travelers who cannot buy full-price tickets in the last minute
  - It happens when demand exceeds inventory
  - Expected lost-sales = 25 – 22 = 3

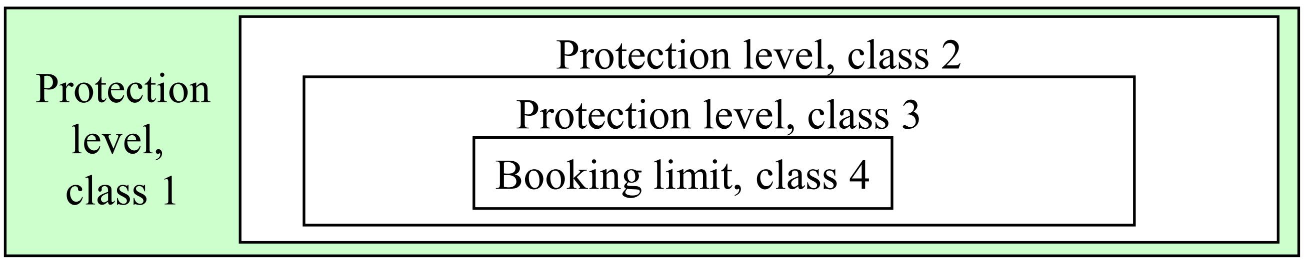
- Expected revenue: includes both high-fare and low-fare
  - Out of the 335 seats, 23 are booked for business travelers, 22 are expected to be sold
  - 312 are for leisure travelers, which can always be sold
  - Expected revenue = $7950 \times 22 + 5250 \times 312 = 1.81$ million

Parallel with Newsvendor Model

	RM with Capacity controls: Two-class strategy	**Newsvendor for Inventory Management**
Decision	Protection level for high fare	Order quantity
Uncertainty demand	For high fare tickets	For the inventory
Overstocking cost	Discounted fare price	Purchase cost minus salvage value
Understocking cost	Full fare price minus discounted fare	Sell price minus purchase cost

Generalization to Multiple Fare Classes

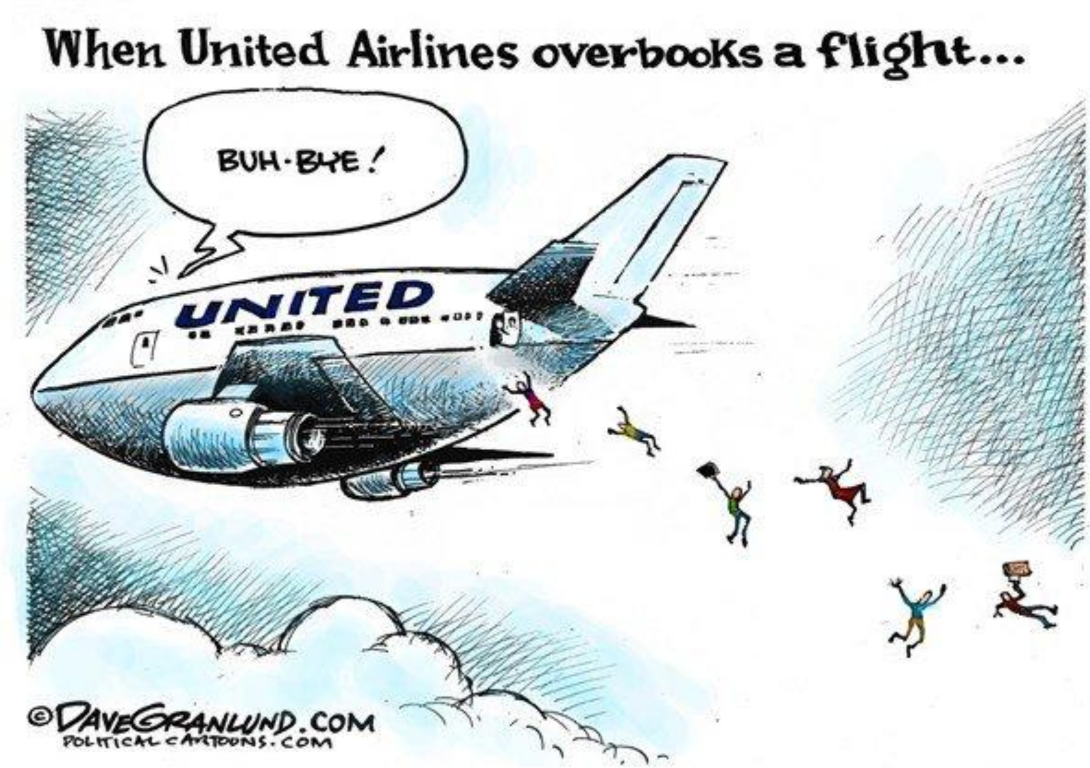
- There can be more than two fare classes, with $f_1 > f_2 > f_3 > \dots > f_n$
- Previous solution can be extended to multiple fare classes, using the nested structure **(Not required for this course!)**



  - Class 1 vs. Classes 2-4: Determine protection level Q_1 for class 1
  - Class 2 vs. Classes 3-4: Determine protection level Q_2 for class 2
  - Class 3 vs. Class 4: Determine protection level Q_3 for class 3

- Solve from high to low iteratively: use newsvendor model in each step (nested booking limit)
- Expected Marginal Seat Revenue Model (EMSR)

Application 2: Overbooking



Overbooking is a common practice in airline industry

- probability that all passengers checking in is smaller than 0.0001

Video: Overbooking in Flights

Video

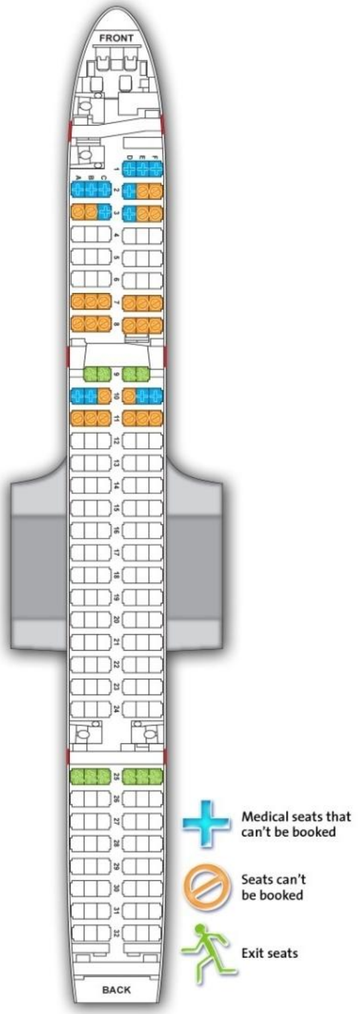
Treatment of Overbooked Passengers

- Passengers are "bumped" if there are not enough seats when they show up
  - Volunteers or involuntary denied boarding

- Costs of bumping
  - Direct cost for the compensation/travel arrangement
  - Implicit cost for ill-will of customers

Overbooking Problem

- Suppose there are 100 seats on a flight from Hong Kong to Singapore
- The number of people who book tickets but do not show up follows a normal distribution with mean of 20 and standard deviation of 10
- Air ticket price = $105
- Total cost of denied boarding: $405
  - Arrangement for travel on another airline: $200
  - Free air ticket: $105
  - Ill-will cost: $100
- How many reservations should the airline take to maximize expected revenue?



Overbooking: Examples

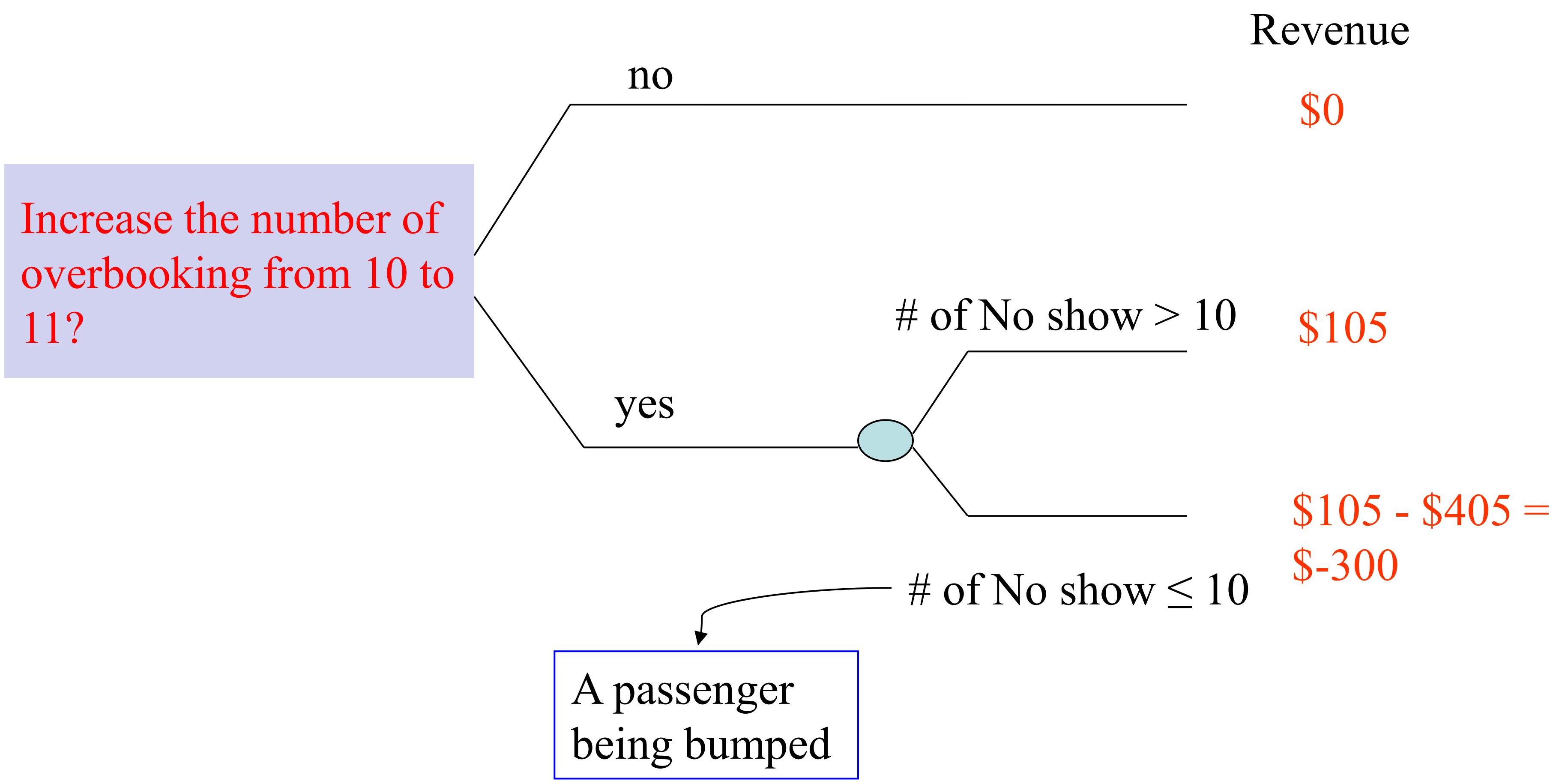
- Suppose you have 100 seats but you overbook 12 seats, i.e., you sell in total 112 seats to the customers

- If the no-show number turns out to be 8, then in total 112 -8 = 104 customers will arrive to gate
- You need to "bump" 4 customers; there will be no empty seats on the flight

- If the no-show number turns out to be 15, then in total 112 – 15 = 97 customers will arrive to gate
- You do not need to bump anyone; there will be 100 -97 = 3 empty seats on the flight

Optimal Overbooking Level

- Use marginal analysis to solve the optimal overbooking level



Convert to a Newsvendor Problem

- In the overbooking problem, the "demand" is the number of people who do not show up
- The "inventory" is the number of overbooked tickets
- Then, the overstocking cost is the penalty when you overbook one more seat but the customer eventually showed up
  - It means you need to bump one more customer
- The understocking cost is the penalty when you overbook one fewer seat but the customer does not show up
  - It means you lose the opportunity to sell one more ticket for no cost

Reframing into Newsvendor Model

- Overstocking cost (too many seats overbooked, bumping a passenger): C_o = \$405 - \$105 = \$300
- Understocking cost (too little seats overbooked, losing a ticket fare): C_u = \$105

- Optimal level of overbooking X^* satisfies

$$\Pr(\#\text{ of no shows} \leq X) = \frac{C_u}{C_u + C_o} = \frac{105}{105+300} = 0.259$$

- Since number of no shows is N(20, 10), we have

$$X^* = 20 + 10z^* = 13.5 \approx 14$$

with $z^* = -0.65$ from distribution table

Service Level in Overbooking Problem

- Number of "bumped" customer
  - # of bumped customer = # of overbook − # of no shows
  - set to zero if # of overbook is smaller than # of no shows

- No customer bumping: number of no shows is larger or equal to number of overbooking

- Probability of bumping = Pr(# of no shows ≤ # of overbook -1)
  - This accounts for the fact that # of no shows must be integers
  - The probability can be computed by z-score

Overbooking Problem: Performance Metrics

- To understand the performance metrics in overbooking problem, let us consider the following example
- Assume the flight has 100 seats

	Scenario 1	Scenario 2
Overbook (Inventory Q)	15	15
No-show (Demand D)	20	10
Total tickets	115	115
Total arrival	95	105
Bumped customers	0	5
Additional tickets sold	15	10
Empty seats	5	0
Leftover (max{Q-D,0})	0	5
Sale (min{Q,D})	15	10
Lost sales (max{D-Q,0})	5	0

Overbooking Problem: Performance Metrics

- In the overbooking problem, we view overbooked seats as inventory, and no-shows as demand
- Expected leftover: it happens when inventory (overbooked seats) exceeds demand (no show)
  - This is the expected # of bumped customers, which happens when # of overbook is larger than # of no-shows
  - Expected leftover = I(-0.65) ×10 = 0.1554×10 = 1.554 ≈ 2
- Expected sales: number of additional customers that are served due to overbooking
  - Two scenarios: # of no-show is greater or smaller than # of overbook
  - Expected sales = # of overbooking - expected # of bumped = 14 – 2 =12

Overbooking Problem: Performance Metrics

- Expected lost sales: number of empty seats at the end (# no show - # overbook)
  - Expected lost sales = 20 – 12 = 8

- Expected revenue from overbooking:
  - Expected Revenue = 105 × Expected sales – 300×Expected bumps = 105 × 12 – 300 × 2 = 660
  - Earn money from the additional tickets sold, lose money from the bumped customers

Key Ingredients for Newsvendor Problem

- Always remember to figure out what is the inventory and what is the demand…
- Two principles: (1) a higher demand benefits our revenue; (2) we aim to use inventory to match the demand
- Then we can identify the overstocking cost (inventory is too high) and the understocking cost (inventory is too low)
- We can also get the performance metrics (leftover, sales, lost-sales) according to their definitions…
  - And interpret them in our specific setting

Parallel with Newsvendor Model

	RM with Capacity controls: Two-class strategy	RM with Capacity controls: Overbooking	Newsvendor for Inventory Management
Decision	Protection level for high fare	Overbooking level	Order quantity
Uncertainty demand	For high fare tickets	Number of no shows	For the inventory
Overstocking cost	Discounted fare price	Bump cost minus fare price	Purchase cost minus salvage value
Understocking cost	Full fare price minus discounted fare	Fare price	Sell price minus purchase cost

Knowledge Points

- Overview of revenue management and customer segmentation
- Capacity-based management: two-class allocation problem and overbooking problem
    - Model formulation
    - Quantitative steps
- Solving the problems by newsvendor model
    - Identification of overstocking and understocking costs
    - Solving optimal protection level
    - Performance metrics (e.g., expected profit and service level)

ISOM 2700: Operations Management

Session 19. Revenue Management II: Price-based RM

Yiwen Shen
Dept. of ISOM, HKUST
Fall 2025

Agenda

- **Overview of price-based revenue management**
- Price optimization based on willingness-to-pay
- Price optimization based on demand model
    - Single- and multi-product pricing
- Other topics for price-based revenue management

What is Revenue Management?

- Revenue Management (RM) was first developed in airline industry in 1980
- It is the practice of maximizing revenue by optimizing product availability (capacity) and price
- Evolution of RM:

Capacity-based RM	Price-based RM
• Widely used in airlines, hotels, and car rental industries • Fixed and perishable resources • …	• Markdown pricing in retailing • Dynamic pricing in e-commerce • Surge pricing by Uber • …

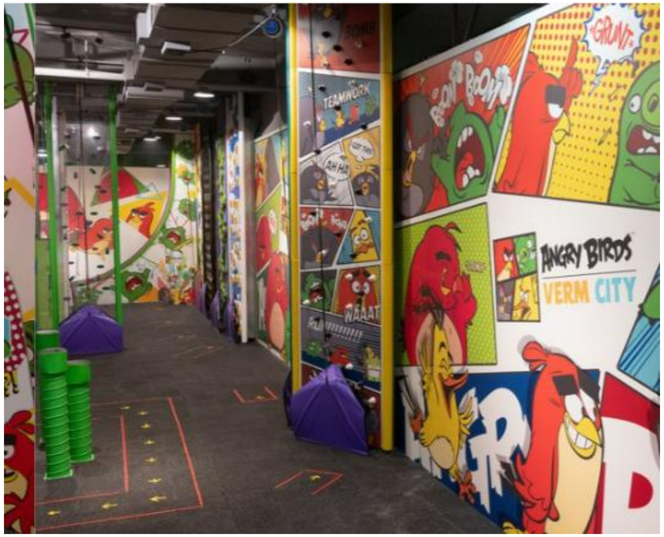
Price Optimization is Everywhere

- Price optimization has become more prevalent in RM, because of:
  - Exploding data for forecasting customer demand
  - Plummeting computing cost
  - More flexibility in changing prices over time

Price Optimization: Verm City in HK

CLIP N CLIMB PARK

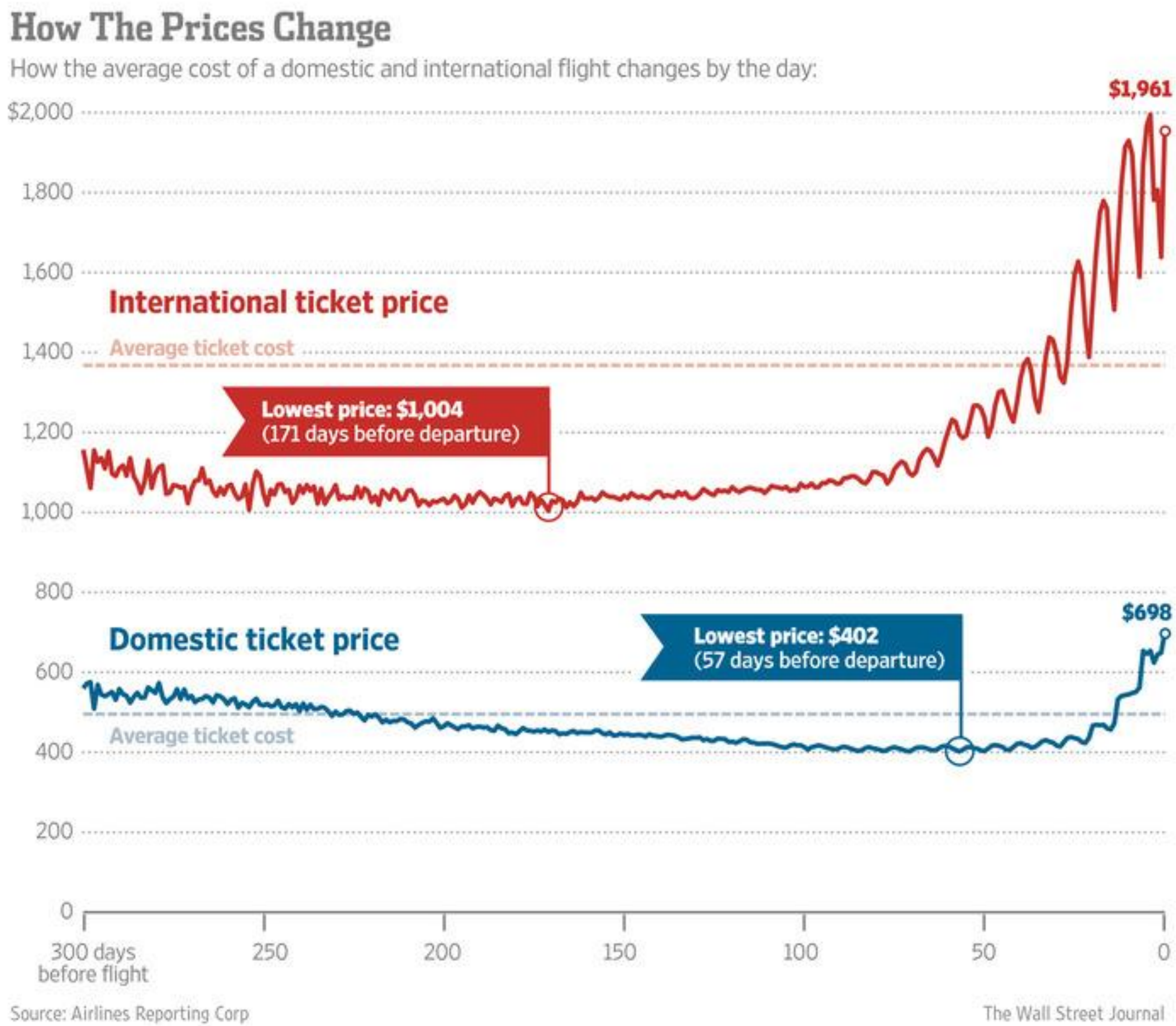
SESSION PASS Single entry for Clip N Climb Park	$250
ANNUAL PASS Unlimited entry for one year	$600
SHOES RENTAL Sneakers	$40



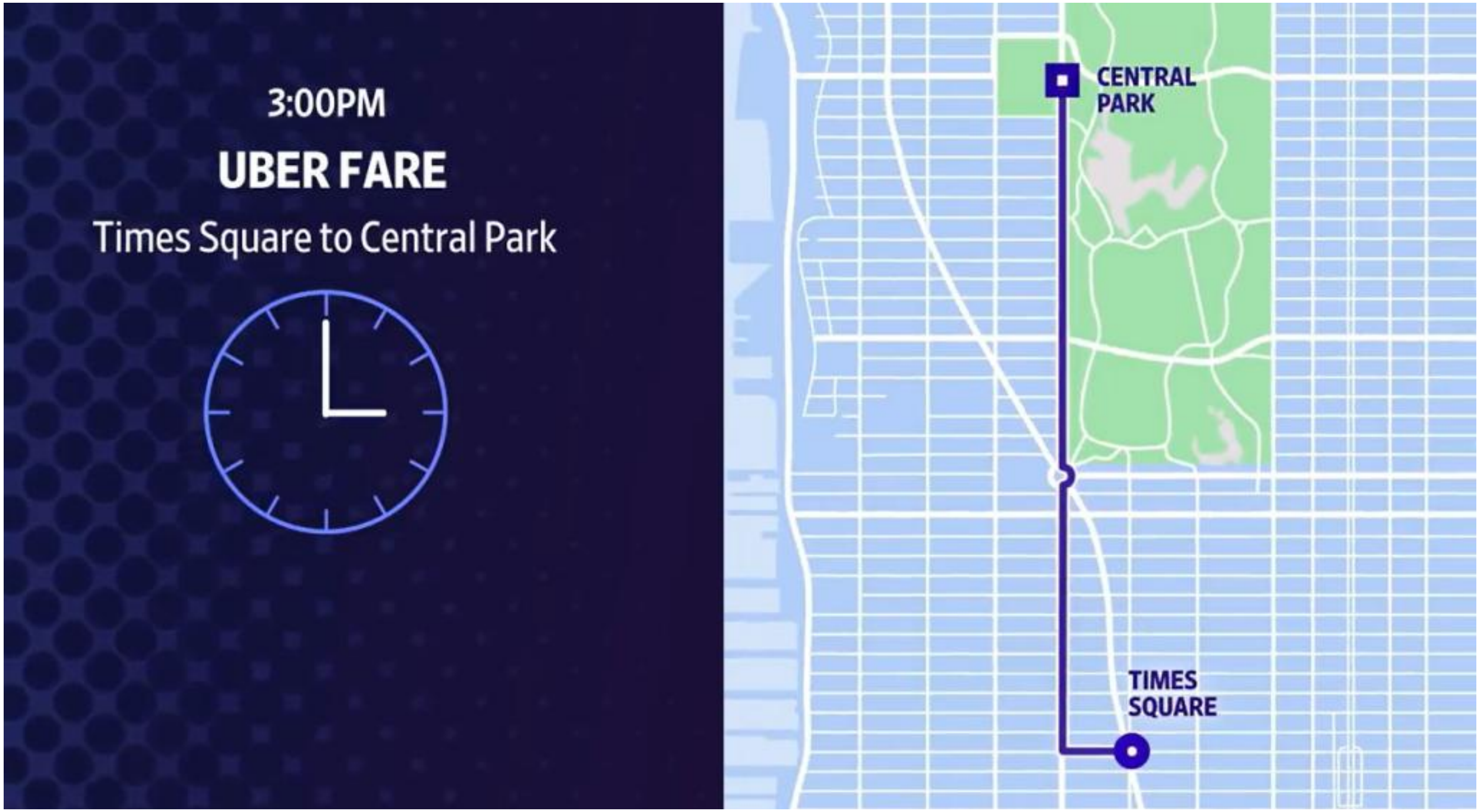
ROCK CLIMBING ZONE

DAY PASS Single day entry for Rock Climbing Zone	$250
1 MONTH MEMBERSHIP Unlimited entry for a month	$800
3 MONTH MEMBERSHIP Unlimited entry for 3 months	$2100
6 MONTH MEMBERSHIP Unlimited entry for 6 months	$3600
1 YEAR MEMBERSHIP Unlimited entry for 1 year	$6000

Examples: Dynamic Pricing



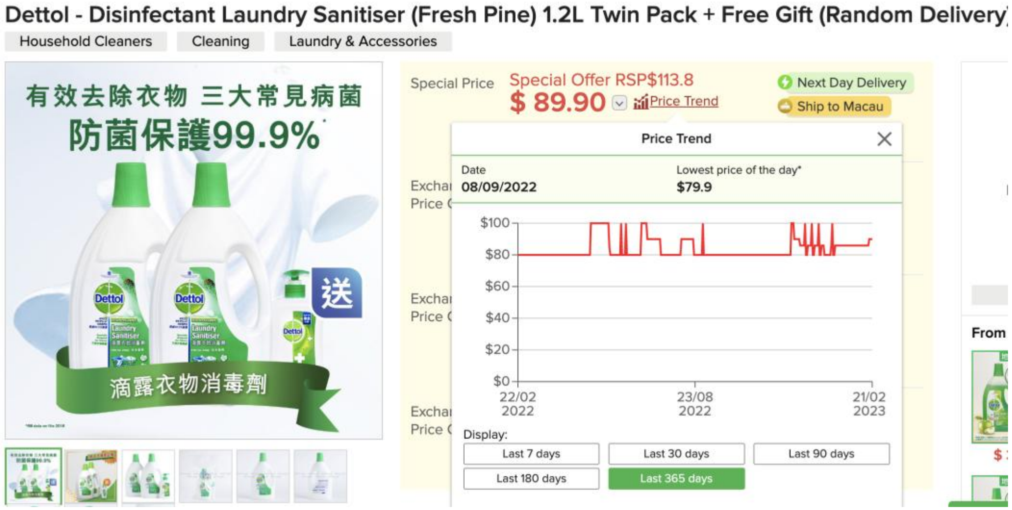
Video: Dynamic Pricing



[Video]

Dynamic Pricing in E-Commerce

HKTV Mall

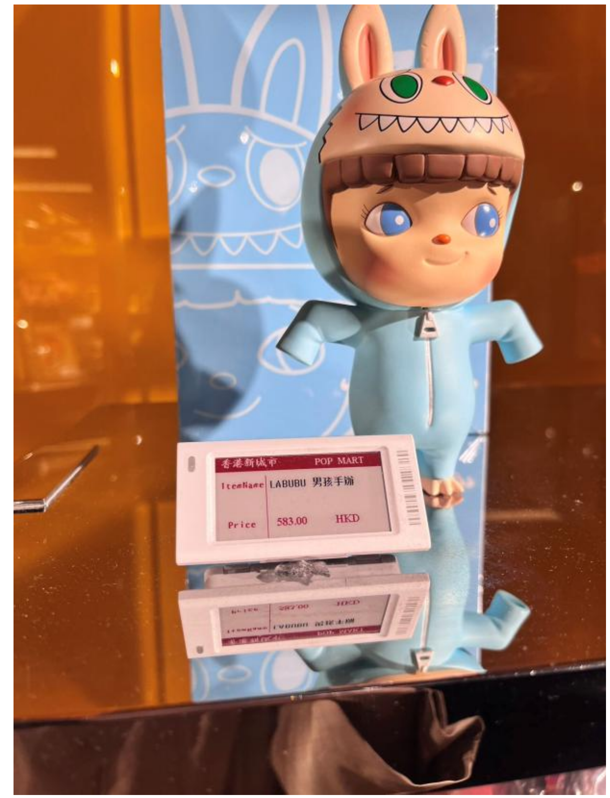


Amazon

About 4% of Amazon's annual revenue is attributable to dynamic pricing!

Dynamic Pricing: Electronic Price Tags

樓上 HK JEBN



Pop Mart

Agenda

- Overview of price-based revenue management
- **Price optimization based on willingness-to-pay**
- Price optimization based on demand model
  - Single- and multi-product pricing
- Other topics for price-based revenue management

Willingness-to-pay

- Willingness-to-pay (WTP) is the maximum cost a customer is willing to pay to buy a unit of product or service

- Customers may have different WTP for the same product

- Firms can segment customers by their WTPs:
  - Airlines charge different prices for leisure and business travelers
  - Book publishers sell student editions at discount price

- WTP is determined by:
  - Customer characteristics: e.g., income and gender
  - Product characteristics: e.g., quality and brand
  - Availability of substitute options

Revenue Maximization based on WTP

- Pricing a textbook

WTP	Frequency
80	10
70	12
60	14
50	28
40	20
30	10
20	6

- How should we set the price to maximize revenue
- Trade-off: higher demand versus higher price

Revenue Maximization based on WTP

- Pricing a textbook

WTP	Frequency	Demand if price = WTP
80	10	10
70	12	22
60	14	36
50	28	64
40	20	84
30	10	94
20	6	100

Check candidate price at each WTP

- Sort WTP from high to low
- Demand at each WTP as cumulative sum of frequency
- Revenue as the product of price and demand

Revenue Maximization based on WTP

- Pricing a textbook

WTP (Price)	Frequency	Demand if price = WTP	Revenue (Price × Demand)
80	10	10	800
70	12	22	1540
60	14	36	2160
50	28	64	3200
40	20	84	3360
30	10	94	2820
20	6	100	2000

Revenue maximized at price = $40

Agenda

- Overview of price-based revenue management
- Price optimization based on willingness-to-pay
- **Price optimization based on demand model**
  - **Single- and multi-product pricing**
- Other topics for price-based revenue management

Single-Product: Linear Demand

- Suppose the demand is a linear function of price p:

$$D(p) = a - bp$$

where *a* is a constant, *b* is the sensitivity to price

- The revenue at price *p* is given by

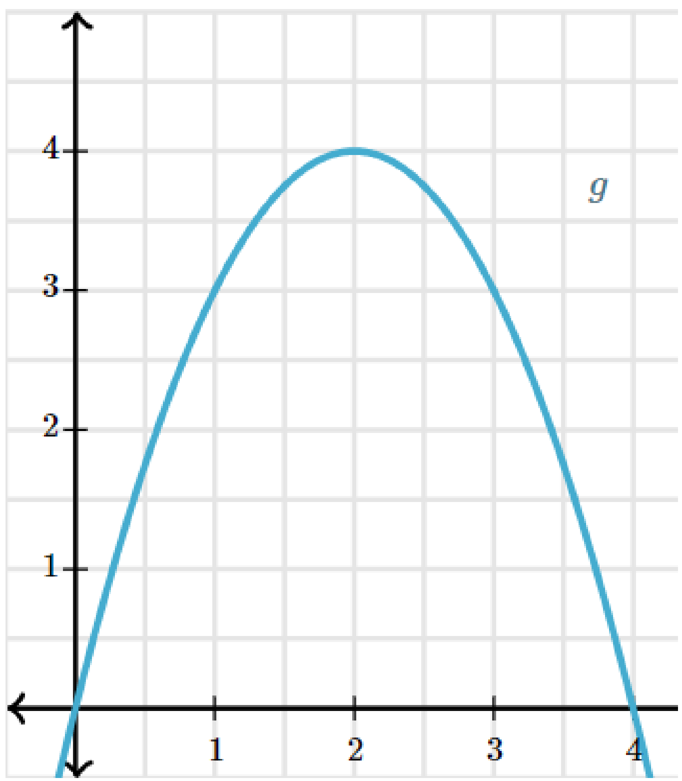
$$R(p) = pD(p) = p(a-bp)$$

- We can solve the revenue-maximizing price p^* as

$$p^* = \frac{a}{2b}$$

Maximizing a Concave Quadratic Function

- For a quadratic function $f(x) = -bx^2 + ax + c$
- When b>0, it is maximized at $x^* = \frac{a}{2b}$
- This is what you need to know for this course!



Single-Product: Example

- Consider a store selling T-shirt. The demand for the T-shirt at price p is given by

$$D(p) = 120 - 20p$$

What is the optimal price and the optimal revenue?

- The optimal price is given by

$$p^* = \frac{120}{40} = 3$$

- The demand at this price is 120-20*3 = 60 units
- The optimal revenue = p*D(p) = 3*60 = 180 dollars

Linear Demand Approximation

- Linear demand function is an efficient way to approximate real-world demand

- It can be estimated by quantitative methods (e.g., linear regression) or qualitative methods (e.g., market survey)

- Advantages of linear demand function:
  - Easy to model, estimate, and communicate
  - Only need limited amount of data

- Disadvantages of linear demand function:
  - Cannot capture non-linear demand-price relationship
  - E.g., marginal price impact may change with price level

Multi-product Pricing

- Consider a cafe store and a tea store that run on a same street
- Denote the coffee price by p_1 and tea price by p_2
- Based on previous sales, the daily demands of the two products are estimated by

$$\text{Coffee: } D_1(p_1, p_2) = 100 - p_1 + 0.5p_2$$

$$\text{Tea: } D_2(p_1, p_2) = 120 + 0.5p_1 - p_2$$

- Note that the demand for each product depends on both prices
- Are tea and coffee substitutes or complements?

Individual Pricing: Solution

- Suppose the tea price is given by p_2, what is the optimal price for the coffee?

- Revenue for coffee store:

$$\begin{aligned} R_1(p_1, p_2) &= p_1 D_1(p_1, p_2) \\ &= -p_1^2 + (100 + 0.5p_2)p_1 \end{aligned}$$

- View p_2 as given, the optimal p_1 is given by

$$p_1^* = 50 + \frac{p_2}{4}$$

- Likewise, suppose the coffee price is given by p_1, the optimal tea price for maximizing the tea store revenue is

$$p_2^* = 60 + \frac{p_1}{4}$$

Equilibrium Prices

- We see that the optimal price of each player depends on the price of the other

- Thus, if one player changes its price, the other player will change as well

- This will eventually lead to an equilibrium, where both players are using their optimal prices

- That is, both players do *not* want to change their price, given the price level of the other one

Equilibrium with Individual Pricing

- At the equilibrium, both players are using their optimal prices
- This requires the following equation system to hold:

$$p_1^* = 50 + \frac{p_2^*}{4}$$
$$p_2^* = 60 + \frac{p_1^*}{4}$$

- Combining these two, we get $(p_1^*, p_2^*) = (69.3, 77.3)$

- Total revenue = $R_1(p_1^*, p_2^*) + R_2(p_1^*, p_2^*) = \$10{,}785$

Pricing the Products Jointly

- Now, suppose the two stores merge into one, what should be the optimal prices of coffee and tea?

- For the merged store, it only cares about the total revenue from both coffee and tea

- It can be calculated as

$$R_1(p_1, p_2) + R_2(p_1, p_2)$$
$$= -p_1^2 - p_2^2 + p_1 p_2 + 100p_1 + 120p_2$$

- We choose the optimal p1 and p2 to maximize the above

Joint Pricing: Solution

- First, we solve the optimal p_1 by viewing p_2 as a given constant. We write total revenue as:

$$R_1(p_1, p_2) + R_2(p_1, p_2) = -p_1^2 + p_1(100 + p_2) - p_2^2 + 120p_2$$

- Using the formula, we have: $p_1^* = 50 + \frac{p_2}{2}$

- Then, we solve the optimal p_2 by viewing p_1 as a given constant. We write total revenue as:

$$R_1(p_1, p_2) + R_2(p_1, p_2) = -p_2^2 + p_2(120 + p_1) - p_1^2 + 100p_1$$

- Using the formula, we have: $p_2^* = 60 + \frac{p_1}{2}$

Joint Pricing: Solution

- We combine the two equations

$$p_1^* = 50 + \frac{p_2^*}{2}$$
$$p_2^* = 60 + \frac{p_1^*}{2}$$

- Combining these two, we get $(p_1^*, p_2^*) = (106.7{,}113.3)$

- Total revenue = $R_1(p_1^*, p_2^*) + R_2(p_1^*, p_2^*) = \$12{,}133.3$

Comparison of Individual and Joint Pricing

	Individual pricing	Joint pricing
p_1^*	69.3	106.7
p_2^*	77.3	113.3
$R_1(p_1^*, p_2^*) + R_2(p_1^*, p_2^*)$	10785	12133.3

- With joint pricing, the total revenue must become larger: benefit of coordination between stores
- With coordination, the prices are higher: coordination reduces the price competition among the stores, thus they both rise prices

Benefit of Pricing Jointly

- Pricing the two products jointly increases the total revenue compared with pricing them individually

- Intuition: the demand for each product depends on both its own price and the price of the other product
  - Interaction between the demands of the two products

- Joint pricing accounts for the impact of each product's price on the other product

Distribution of Total Revenue

	Individual pricing	Joint pricing
$R_1(p_1^*, p_2^*)$ $+ R_2(p_1^*, p_2^*)$	10785	12133.3
$R_1(p_1^*, p_2^*)$	4806	5329.7
$R_2(p_1^*, p_2^*)$	5979	6803.7

- With joint pricing, both stores are better off in this example. So, they should be happy to change
- This does NOT always hold. It is possible that one player's revenue becomes lower in the joint pricing
- The distribution of revenue is an important issue, and we will discuss this in later sessions

Distribution of Total Revenue

- Consider the following example with two firms

Revenue	Separate pricing	Joint pricing
Firm 1	1500	1400
Firm 2	1200	1600
Firm 1 + Firm 2	2700	3000

- Using joint pricing indeed improves the total revenue of the two firms: from 2700 to 3000
- However, Firm 1 may not be willing to switch to joint pricing because it now gets a lower revenue: from 1500 to 1400
- How can you handle this to improve the outcomes for both?
- Solution: better distribution of revenue!
- For example, let Firm 2 transfers 300 to Firm 1 under joint pricing

Pricing for Two Products

- Pricing for two products with linear demands:

$$D_1 = a_1 - b_{11}p_1 - b_{12}p_2 \text{ and } D_2 = a_2 - b_{21}p_1 - b_{22}p_2$$

- Here p_1 and p_2 are prices; a_1, b_{11}, b_{12}, a_2, b_{21}, b_{22} are constant coefficients given by the problem

	Pricing Individually	Pricing Jointly
Objective	Set p_1 to maximize R_1 Set p_2 to maximize R_2	Set p_1 and p_2 to maximize $R_1 + R_2$
Equation systems	$2b_{11}p_1 + b_{12}p_2 = a_1$ $b_{21}p_1 + 2b_{22}p_2 = a_2$	$2b_{11}p_1 + (b_{12}+b_{21})p_2 = a_1$ $(b_{12}+b_{21})p_1 + 2b_{22}p_2 = a_2$

Pricing for Two Products: Example

- Back to the tea and coffee example, we have

$$D_1(p_1, p_2) = 100 - p_1 + 0.5p_2$$
$$D_2(p_1, p_2) = 120 + 0.5p_1 - p_2$$

- To use our formula, we first specify the coefficients as

$$a_1=100,\ b_{11}=1,\ b_{12}=-0.5,\ a_2=120,\ b_{21}=-0.5,\ b_{22}=1$$

- When pricing individually, we can solve prices from

$$2p_1 - 0.5p_2 = 100 \ \ and \ -0.5p_1 + 2p_2 = 120$$

- When pricing jointly, we can solve prices from

$$2p_1 - p_2 = 100 \ \ and \ \ -p_1 + 2p_2 = 120$$

Partial Derivatives

- The maximum of a function is generally obtained when its partial derivatives equal to zero
  - Partial derivative measures the impact of a given variable, keeping other variables fixed

- Suppose a is a constant, we have

$$\partial(ap_1)/\partial p_1 = a \text{ and } \partial(ap_1^2)/\partial p_1 = 2ap_1$$

$$\partial(ap_1p_2)/\partial p_1 = ap_2 \text{ and } \partial(ap_1p_2)/\partial p_2 = ap_1$$

- (NOT Required)

Optimality Conditions

- Individual pricing: set p_1 to maximize $R_1(p_1, p_2)$ and set p_2 to maximize $R_2(p_1, p_2)$
- Optimality conditions:

$$\partial R_1(p_1, p_2)/\partial p_1 = 0$$
$$\partial R_2(p_1, p_2)/\partial p_2 = 0$$

- Joint pricing: set p_1 and p_2 to maximize $R_1(p_1, p_2) + R_2(p_1, p_2)$
- Optimality conditions:

$$\partial(R_1(p_1, p_2) + R_2(p_1, p_2))/\partial p_1 = 0$$
$$\partial(R_1(p_1, p_2) + R_2(p_1, p_2))/\partial p_2 = 0$$

- (NOT Required)

Knowledge Points

- Overview of price-based revenue management
- Price optimization based on WTP
- Price optimization based on demand function: single product
- Price optimization based on demand function: two products
  - Price individually or price jointly
- Other topics: bundling selling, decoy effect, etc

ISOM 2700: Operations Management

Session 20 Managing Supply Chain: Introduction

Yiwen Shen
Dept. of ISOM, HKUST
Fall 2025, HKUST

Pricing for Two Products

- Pricing for two products with linear demands:

$$D_1 = a_1 - b_{11}p_1 - b_{12}p_2 \text{ and } D_2 = a_2 - b_{21}p_1 - b_{22}p_2$$

- Here p_1 and p_2 are prices; a_1, b_{11}, b_{12}, a_2, b_{21}, b_{22} are constant coefficients given by the question

	Pricing Individually	**Pricing Jointly**
Objective	Set p_1 to maximize R_1 Set p_2 to maximize R_2	Set p_1 and p_2 to maximize $R_1 + R_2$
Equation systems	$2b_{11}p_1 + b_{12}p_2 = a_1$ $b_{21}p_1 + 2b_{22}p_2 = a_2$	$2b_{11}p_1 + (b_{12}+b_{21})p_2 = a_1$ $(b_{12}+b_{21})p_1 + 2b_{22}p_2 = a_2$

Pricing for Two Products: Example

- Back to the tea and coffee example, we have

$$D_1(p_1, p_2) = 800 - 2p_1 + 0.5p_2$$
$$D_2(p_1, p_2) = 600 + 0.5p_1 - 2p_2$$

- Remember we want to match the form:

$$D_1 = a_1 - b_{11}p_1 - b_{12}p_2 \text{ and } D_2 = a_2 - b_{21}p_1 - b_{22}p_2$$

- Comparing the coefficients, we have

$$a_1 = 800,\ b_{11} = 2,\ b_{12} = -0.5,$$
$$a_2 = 600,\ b_{21} = -0.5,\ b_{22} = 2$$

Example: Individual Pricing

- What are the optimal prices when we price the two products individually (setting each price to maximize its own revenue)?
- For individual pricing, we use the following equation system:

$$2b_{11}p_1 + b_{12}p_2 = a_1$$
$$b_{21}p_1 + 2b_{22}p_2 = a_2$$

- Plugging the values of a_1, b_{11}, b_{12}, a_2, b_{21}, b_{22}, this becomes

$$4p_1 - 0.5p_2 = 800$$
$$-0.5p_1 + 4p_2 = 600$$

- By them, we can solve the optimal prices as

$$p_1 = 222.2,\ p_2 = 177.8$$

Example: Individual Pricing

- Using optimal prices under individual pricing, what are the demands and revenues for the two products?
- Given the optimal prices:

$$p_1 = 222.2, p_2 = 177.8$$

- Plug them in, the demands are

$$D_1(p_1, p_2) = 800 - 2p_1 + 0.5p_2 = 444.5$$
$$D_2(p_1, p_2) = 600 + 0.5p_1 - 2p_2 = 355.5$$

- The optimal revenues are given by

$$R_1(p_1, p_2) = p_1 \times D_1(p_1, p_2) = 222.2 \times 444.5 = 9.88 \times 10^4$$
$$R_2(p_1, p_2) = p_2 \times D_2(p_1, p_2) = 177.8 \times 355.5 = 6.32 \times 10^4$$

Example: Joint Pricing

- What are the optimal prices when we price the two products jointly (setting the two prices to maximize the total revenue)?
- For joint pricing, we use the following equation system:

$$2b_{11}p_1 + (b_{12}+b_{21})p_2 = a_1$$
$$(b_{12}+b_{21})p_1 + 2b_{22}p_2 = a_2$$

- Plugging the values of a_1, b_{11}, b_{12}, a_2, b_{21}, b_{22}, this becomes

$$4p_1 - p_2 = 800$$
$$-p_1 + 4p_2 = 600$$

- By them, we can solve the optimal prices as

$$p_1 = 253.3,\ p_2 = 213.3$$

Example: Joint Pricing

- Using optimal prices under joint pricing, what are the demands and revenue for the two products?
- Given the optimal prices:

$$p_1 = 253.3, p_2 = 213.3$$

- Plug them in, the demands are

$$D_1(p_1, p_2) = 800 - 2p_1 + 0.5p_2 = 400$$
$$D_2(p_1, p_2) = 600 + 0.5p_1 - 2p_2 = 300$$

- The optimal revenues are given by

$$R_1(p_1, p_2) = p_1 \times D_1(p_1, p_2) = 253.3 \times 400 = 1.01 \times 10^5$$
$$R_2(p_1, p_2) = p_2 \times D_2(p_1, p_2) = 213.3 \times 300 = 6.40 \times 10^4$$

Agenda

- Overview of price-based revenue management
- Price optimization based on willingness-to-pay
- Price optimization based on demand model
  - Single- and multi-product pricing
- **Other topics for price-based revenue management**

Markdown Pricing

- Markdown is a permanent reduction in price whereas promotions are temporary

Markdowns: Reasons and Concern

- Obsolescence and spoilage of inventory
  - E.g., release of the next generation of iPad
- Time of use: winter coats in February, costumes for Halloween, Christmas trees
- Deterioration: one-day old bread sold at half prices
- Cannibalization effect: customers will intentionally wait for the markdown, may erode regular sales

Bundle Pricing

- A business strategy where companies group several products together into a bundle
    - …and sell the bundle at a single price
    - rather than attribute individual prices to each item.

- Buyers and sellers can simultaneously benefit from bundling

- Widely used by restaurants, retail stores, and internet companies (e.g., gaming, entertainments)

- Intuition: exploiting the heterogeneity in customers' willing-to-pay for different products (a type of pooling)

Bundle Pricing: Example

- Two cable buyers, “sports lover” and “history lover”

- Willingness-to-pay for the two channels

	ESPN	History channel
Sports lover	$10	$3
History lover	$3	$10

- Two pricing strategies:
  - Each single channel for $9
  - A bundle of two channels for $11

Bundle Pricing: Example

	ESPN	History channel
Sports lover	$10	$3
History lover	$3	$10

- Separate pricing ($9 each): sports lovers would buy ESPN; history lovers would buy history channel
  - Total revenue = $9+$9 = $18
  - Total customer surplus = $1+$1 = $2
- Bundling pricing ($11 for two): both customers will buy two channels
  - Total revenue = $11+$11 = $22
  - Total customer surplus = $2+$2 = $4

Psychological Pricing

- Decoy effect: consumers tend change preference between two options when also presented with a third option
  - Even the third option is asymmetrically dominated (decoy)
  - Subscription options: online, print, online + print

- The Magic Number 9: charm prices ($49, $79, $99, etc) can boost sales compared with to nearby prices
  - The leftmost digit matters the most (e.g., $259, $299, and $329)

Decoy Effect: Example

- Consider a hard-drive store, it offers two products
  - A: price $400, storage 300 GB
  - B: price $300, storage 200 GB
- Some customers prefer A and some prefer B
- Now, add a product C to the option list
  - A: price $400, storage 300 GB
  - B: price $300, storage 200 GB
  - C: price $450, storage 250 GB
- Rational customer never buys C because it is dominated by A
- However, it can lead more customers to chose A
  - A is better than C in both aspects, but B is only partially better than C

Agenda

- **Introduction to Supply Chain Management**

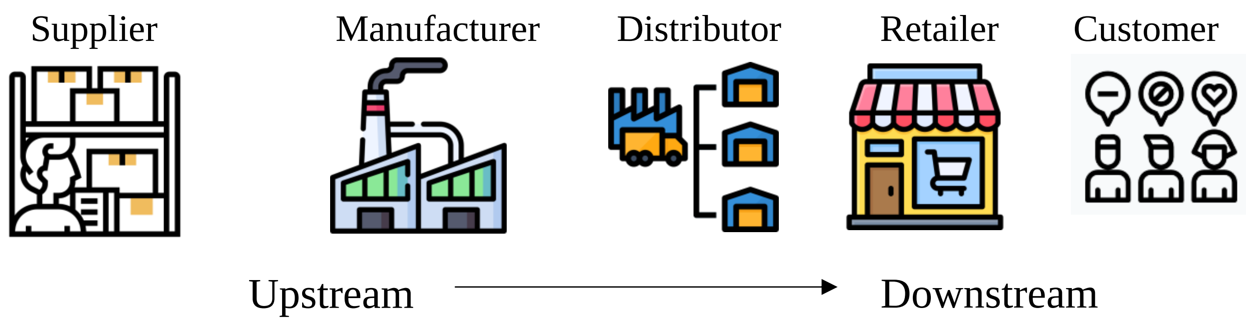
What is a Supply Chain?

Source: TechTarget

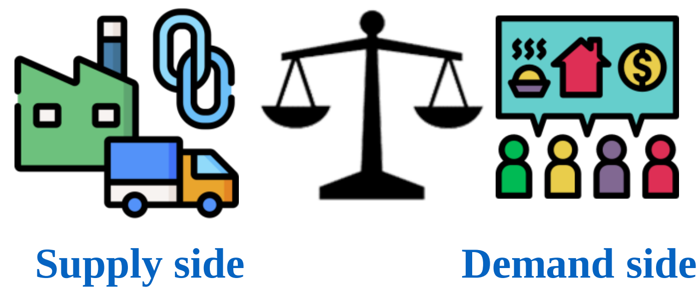
A supply chain is the system of organizations, people, activities, information and resources involved in moving a product or service from supplier to customer

What is a Supply Chain?

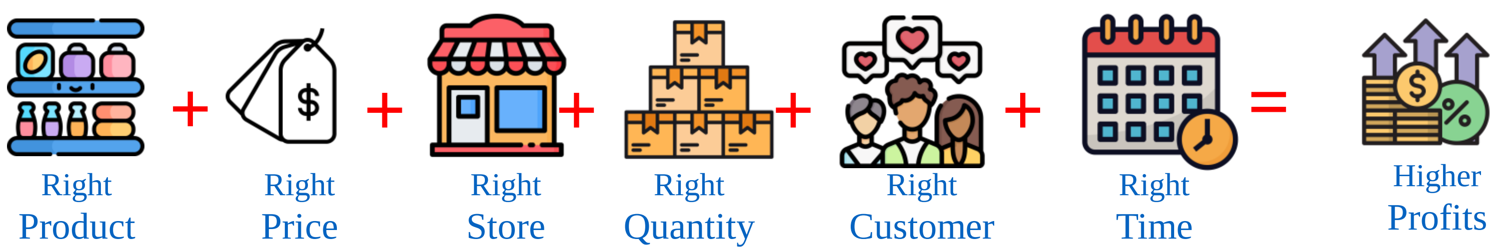
- A supply chain consists of



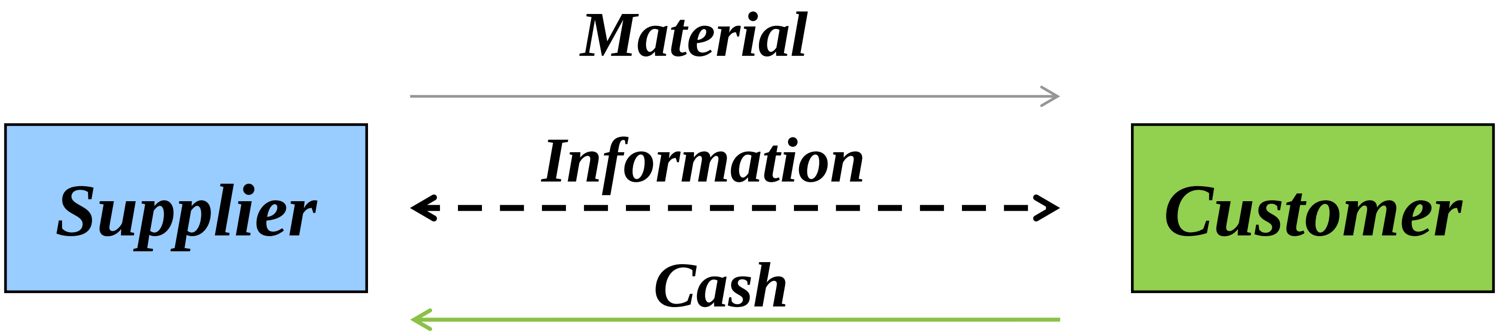
- It aims to Match Supply and Demand profitably for products and services

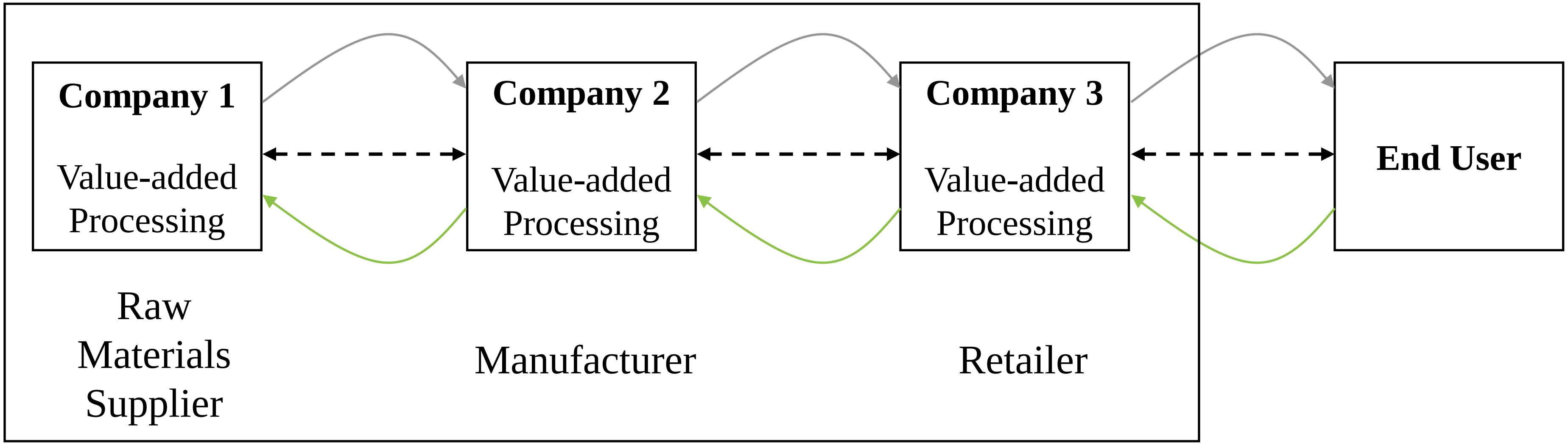


- Increase revenue and reduce cost



Flows in a Supply Chain

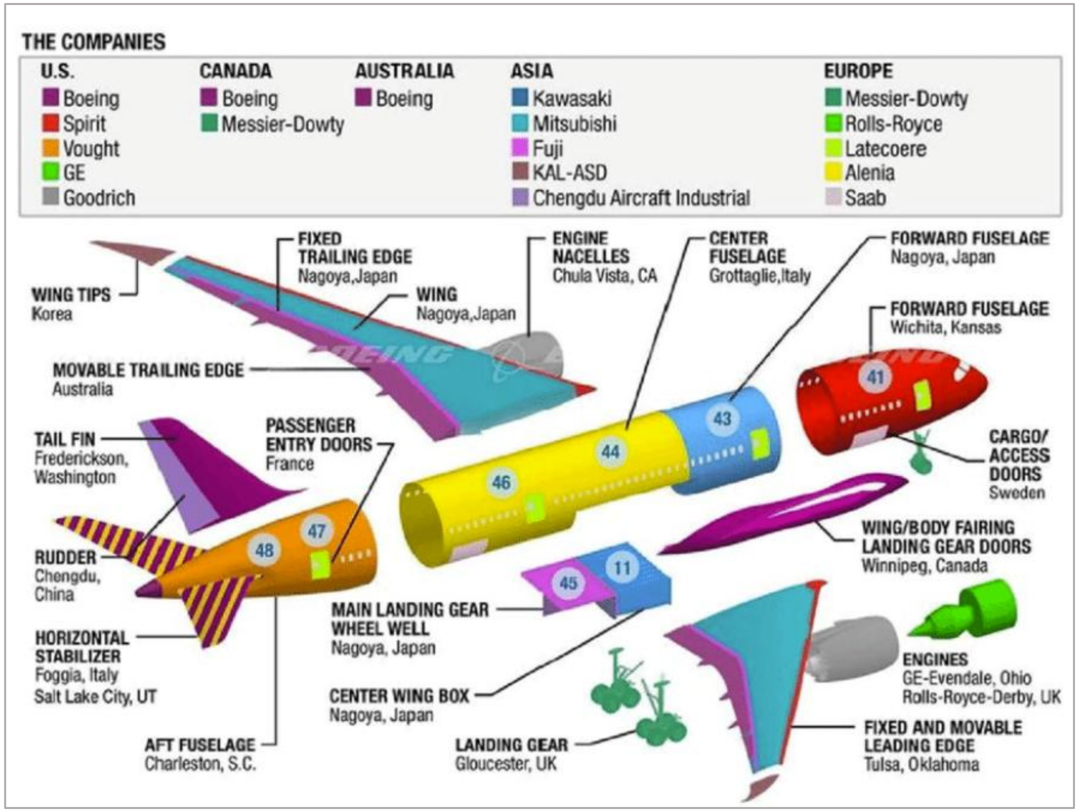




Supply Chain Management: Global Thinking

Supply-chain management (SCM): a total system approach to manage the

- entire flow of information, materials, and services
- from suppliers through other entities to the customer



Global Supply Chain for Boeing

Source

Supply Chain for a US-made Truck

Why Even U.S.-Made Trucks Aren't Safe From Tariffs
Source: The Wall Street Journal

Video

SC Challenges and AAA Strategies

- Increasing uncertainties and rapid changes in demand/supply
  - Uncertainty drives the needs for flexibility
  - Strategy: Agility

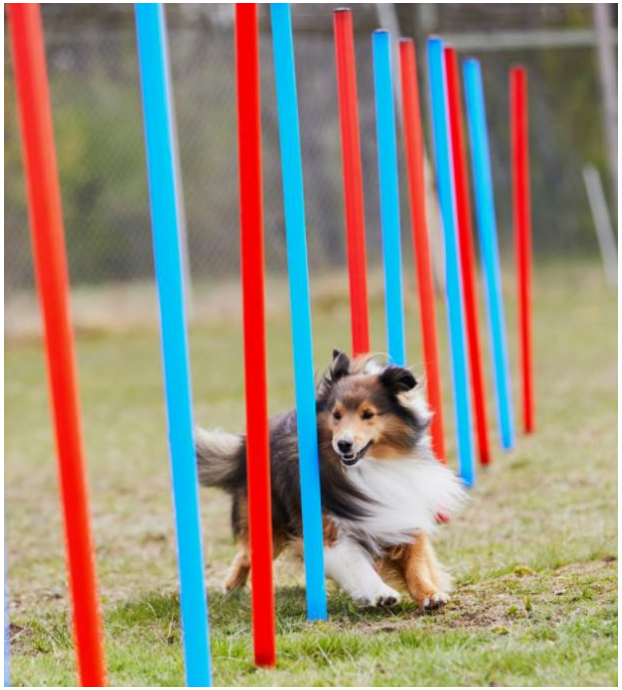
- Shortening product and technology cycles
  - Need for dynamic instead of static supply chains
  - Strategy: Adaptability

- Complex supply networks and many suppliers/customers
  - Need for coordinating interests of multiple players
  - Strategy: Alignment

Building Agility

- Agility: a company's ability to quickly adjust its strategy to meet rapidly changing supply chain requirements
  - e.g., procurement, inventory management, and delivery

- Companies with high agility can navigate change with ease and take advantage of new opportunities



"Sensible" Sense

↓

"Responsive" Response

Building Adaptability

Adaptability: a company's ability to meet structural shifts in markets and supply chain environments

- e.g., changes in technologies, supplies, and demands

Right product
Right market
Right time

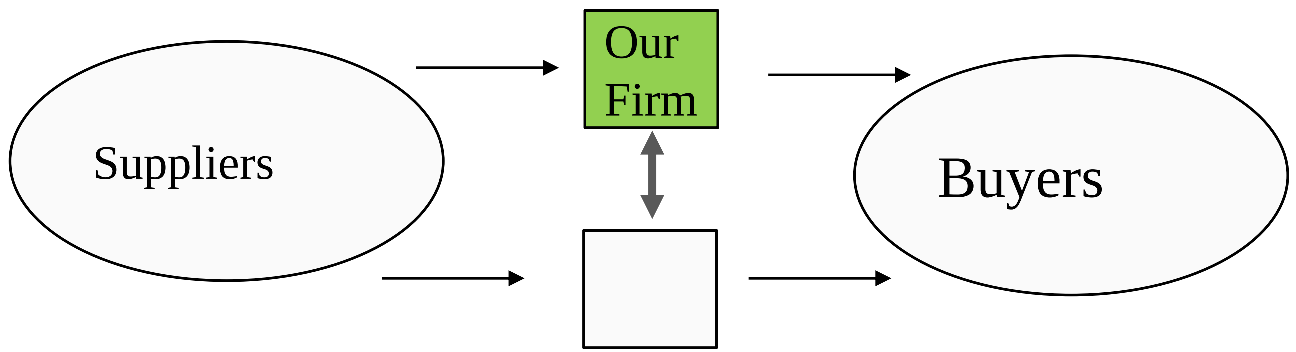
Building Alignment

Alignment: coordination between participants in supply chain to achieve better outcomes for all players

- efficient sharing of information, risks, and rewards
- internal and external alignments

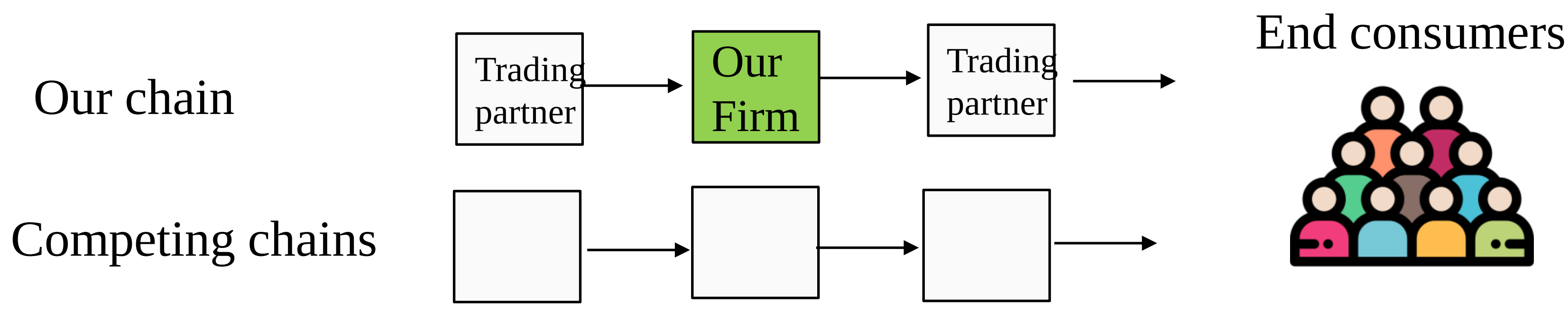
New Perspective from SCM

OLD: "We are a firm competing in our industry ..

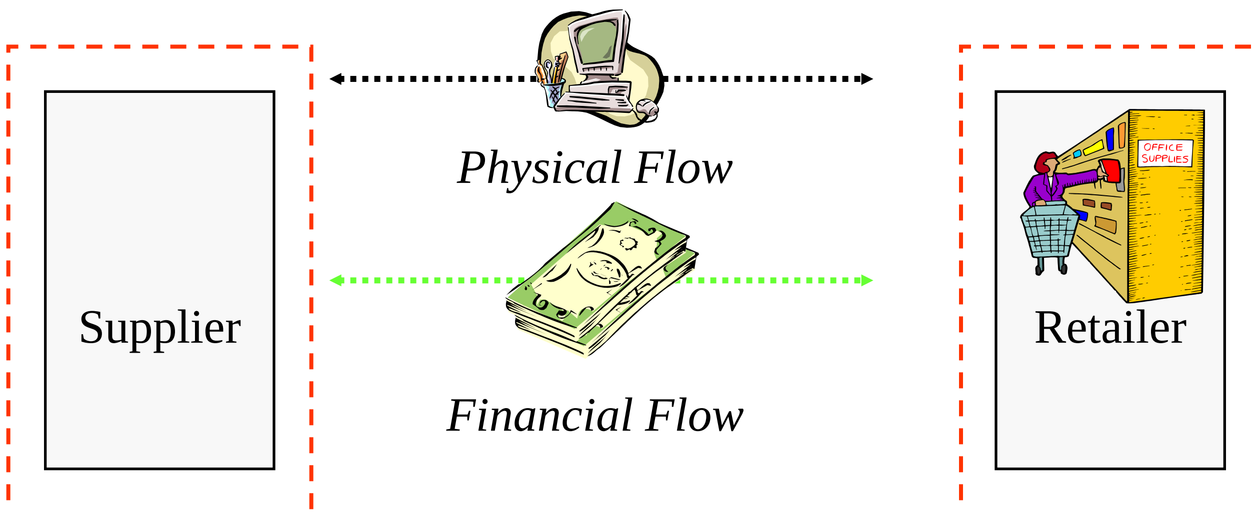


Our industry competition

SCM: "We are part of a supply chain competing with other chains for end consumers."



Coordination in Supply Chain: Contract



- A contract is a set of rules to control / modify the goods/cash flows in the supply chain
- To be effective, the contract should be verifiable and enforceable

Non-alignment: Double Marginalization

- With traditional contract, suboptimal supply chain performance can occur
- This is because each firm makes decisions based on its own profit, not the supply chain's total profit
- This is called double marginalization, reflecting an absence of alignment of interest

Variability in Supply Chain: Example

- Consider a supply chain with a retailer and three-levels of suppliers from downstream to upstream
- Relationship: current-month ordering + start-of-month inventory – current-month demand = end-of-month inventory
- The retailer/suppliers take a simple ordering/production rule:
  - To make the end-of-month inventory equal the current-month downstream demand
- This means we need to set: current-month ordering = 2*current-month demand - start-of-month inventory

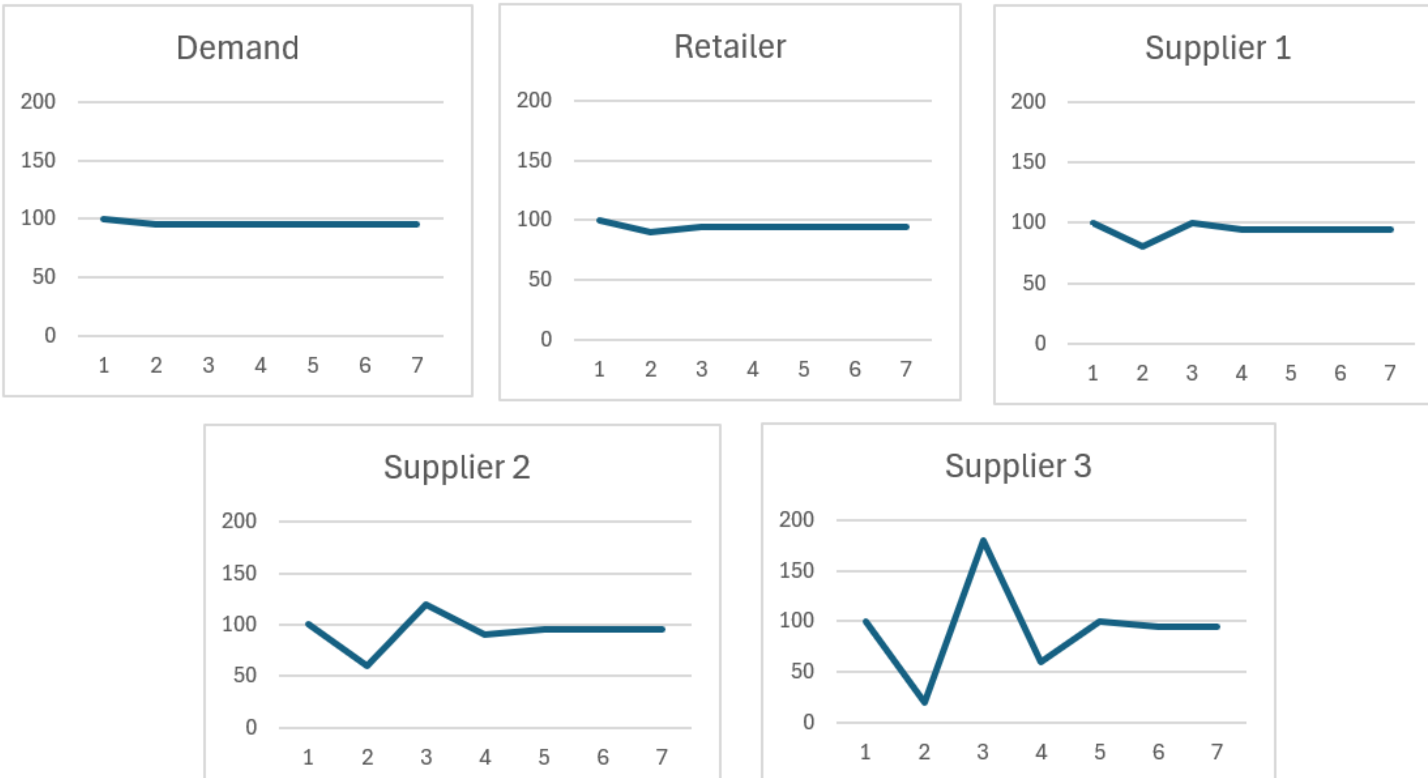
Variability in Supply Chain

- Assume the customer demand drops from 100 to 95, what is the impact on supply chain
- What do you observe in the variability of production quantity?

	Supplier 3		Supplier 2		Supplier 1		Retailer		Customer
	Prodn	Stock	Prodn	Stock	Prodn	Stock	Prodn	Stock	Demand
Month 1	100	100	100	100	100	100	100	100	100
		100		100		100		100	
Month 2	20	100	60	100	80	100	90	100	95
		60		80		90		95	
Month 3	180	60	120	80	100	90	95	95	95
		120		100		95		95	
Month 4	60	120	90	100	95	95	95	95	95
		90		95		95		95	
Month 5	100	90	95	95	95	95	95	95	95
		95		95		95		95	
Month 6	95	95	95	95	95	95	95	95	95
		95		95		95		95	
Month 7	95	95	95	95	95	95	95	95	95
		95		95		95		95	

Upstream → Downstream

Amplification of Variability



- The variability of upstream orders is larger than that of downstream demands
- This is seen in manufacturing, retail, food & beverage, automobiles, etc

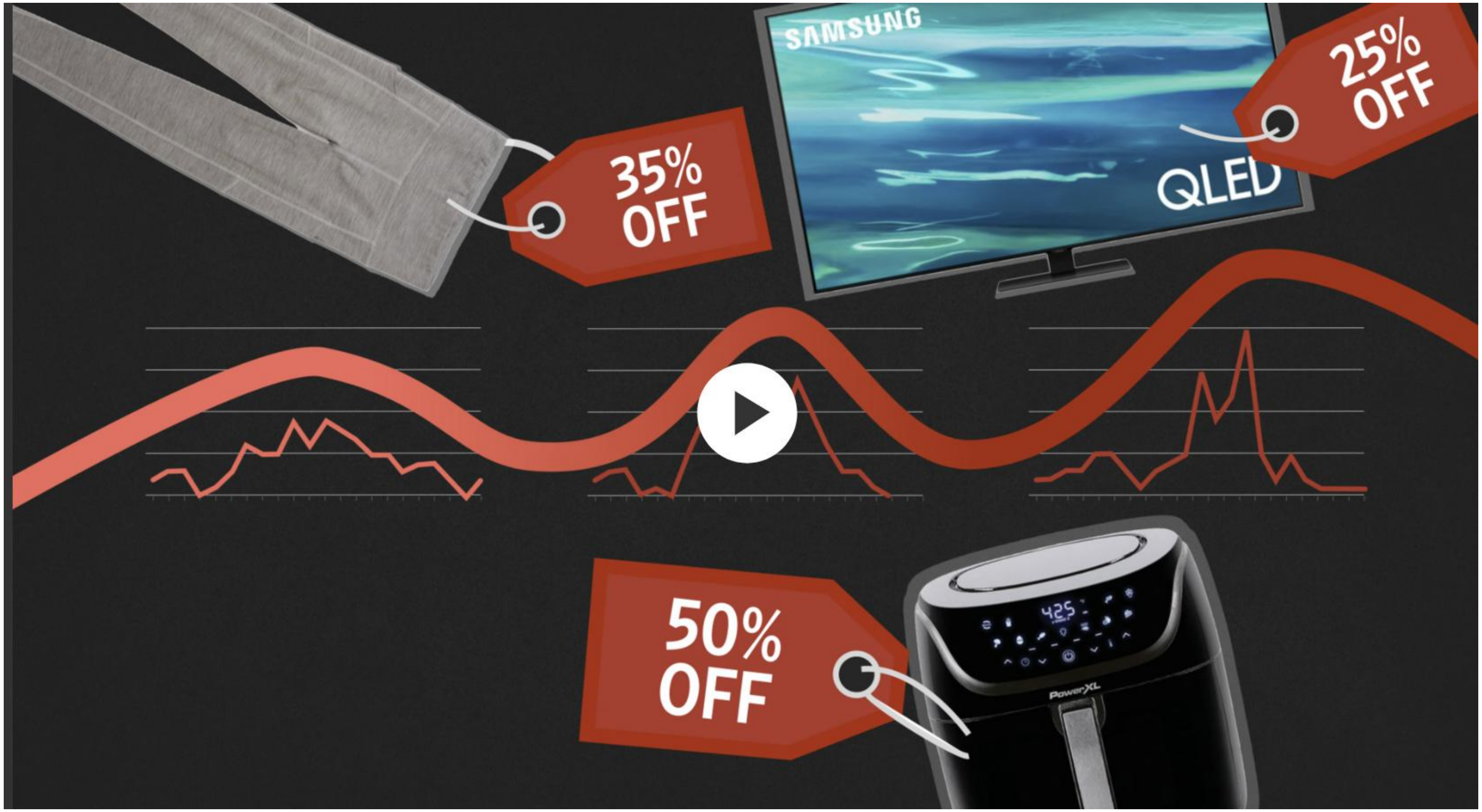
Bullwhip Effect

- Bullwhip Effect: variability of demand at one level is greater than the variability of demand at the next downstream level
  - It can feed on itself when supply chain has multiple layers
  - A common challenge in supply chain management

- Reasons: order synchronization/batching, trade promotion, forward buying, overreactive ordering (not required)

- Solution: collaborative planning, forecasting, and replenishment

Impact of Bullwhip Effect: Video



Why Everything Is On Sale: The Bullwhip Effect
Source: The Wall Street Journal

Video

Discussion: Pandemic and Supply Chain

- The past COVID-19 pandemic has significantly, negatively affected the global supply chain
  - A big disruption in trade, finance, logistics, and workforce
- Shortage for many products (e.g., chips, automobiles), increased delivery time; inflation in US; economic slowdown in China
  - Exposed the vulnerability of supply chain for many firms
- As responses, firms need to put more attention on mitigating and managing hidden risks
  - Diversify supply base, build up safety inventory, process innovation

Reshaping Supply Chain Post Pandemic

Here's How Supply Chains Are Being Reshaped for a New Era of Global Trade

Nearshoring. Automation. Supplier diversification. Sustainability. Companies are adapting their operations to changing market pressures and geopolitics.

Source: The Wall Street Journal (April 2023)

- Single-sourcing to multi-sourcing: less reliance on Asia (particularly China)
- Regionalization: production closer to where companies expect to sell the goods
- Cost minimization (efficiency) vs. risk minimization (robustness)
- Sustainable supply chain (carbon footprint) and automation

Disruptions & Opportunities in Supply Chain

- Recent disruptions in supply chain: pandemic, dockworker strike, hurricane, middle east conflict, trade wars, tariffs
  - Firms need contingency plan, better understanding of supply chain, balancing of multiple objectives
- This brings rich opportunities for risk management in supply chain
  - Data analytics (e.g., import-export data, shipping records)
  - Artificial-intelligence tools for supply chain network

Source: [The Wall Street Journal](#) (Nov 2024)

Video: Reforms in Supply Chain

[Video]

Discussion: Tariff-War and Supply Chain

- From this year, President Trump has started a tariff war that significantly, negatively affected the global supply chain

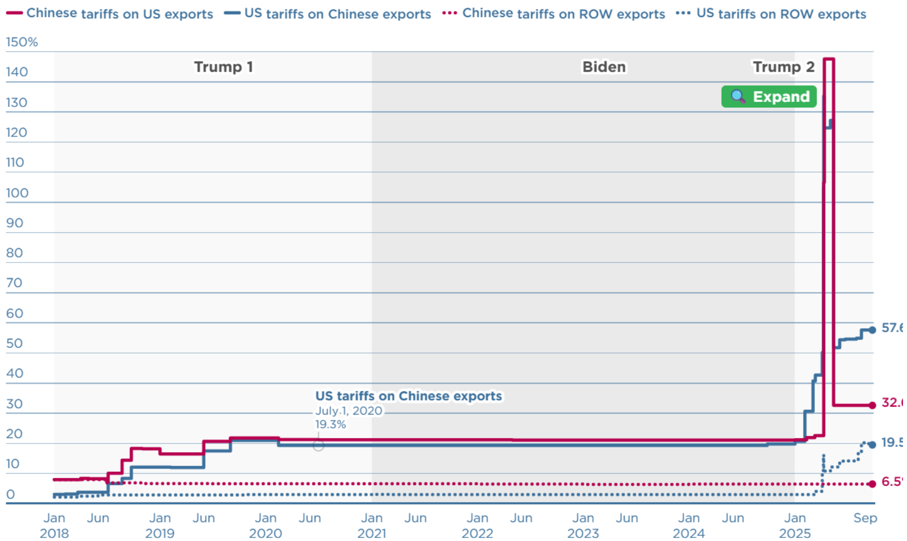
- Very high tariff for MANY countries/regions and product types
  - The "reciprocal" tariff rate is calculated in a strange way
  - Additional tariff paused for most countries except China

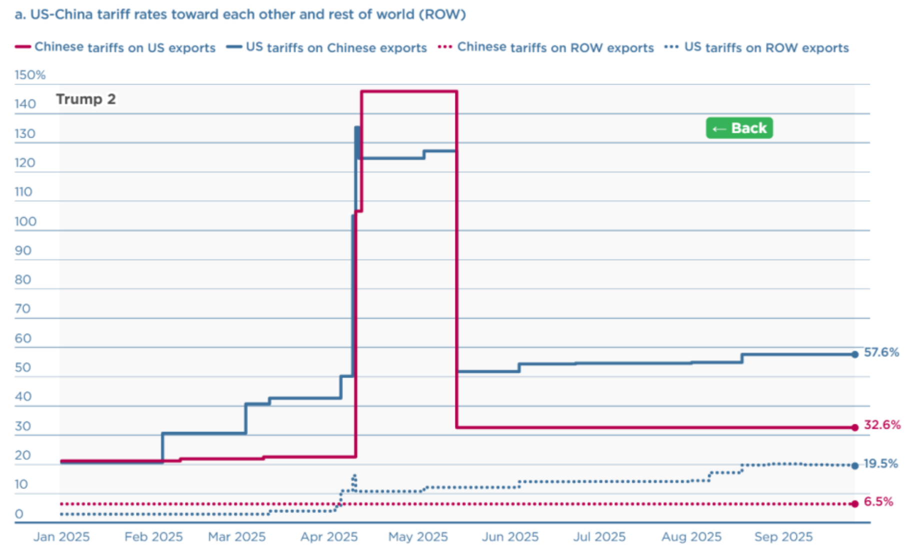
- Super high tariff rate (currently) imposed on China (mainland)

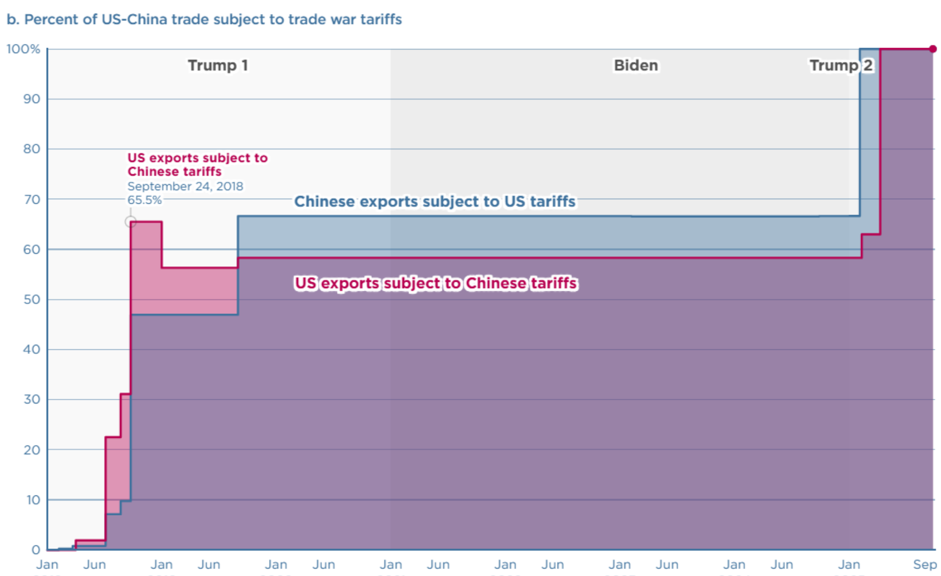
- China has retaliated in a tit-for-tat way with similar tariff rate

- How will this affect the global supply chain in the future?

A (Rough) Timeline for China Tariffs







Source: [PIIE Report]

Also see the summary by JP Morgan at [Link]

Impact on Global Supply Chain and Economics

Source: The Wall Street Journal/MSN (Apr 2025)

- Tariff leads to a disorderly economic decoupling between US and China

- Impact on US: hard to find qualified suppliers, stop of investment, rise in price, more inflation, inventory problem, thinner margin, disturbance in financial market…

- Impact on China: loss of biggest consumer market, fewer demand for products, cancelled orders, layoffs and unemployment, deflationary pressure, rebalancing economy towards consumption…

Tariff Might Not Achieve its Goal

'Derisking' China-Reliant Supply Chains Is Creating New Risks

The U.S. and Chinese economies appear to be growing apart. The real story is more complicated, and worrying.

Source: The Wall Street Journal (Jan 2024)

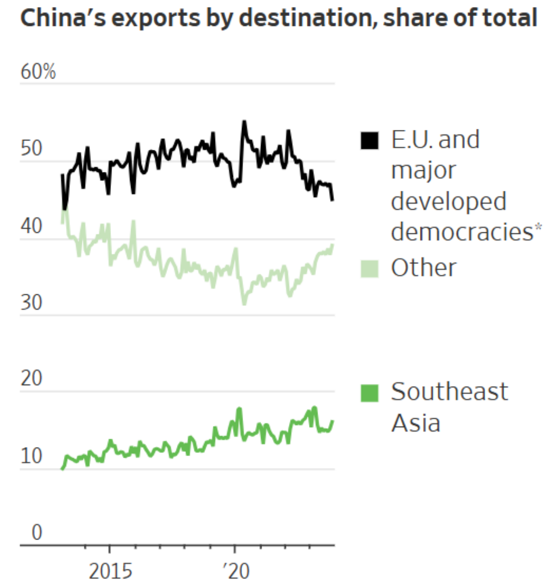
Still coupled

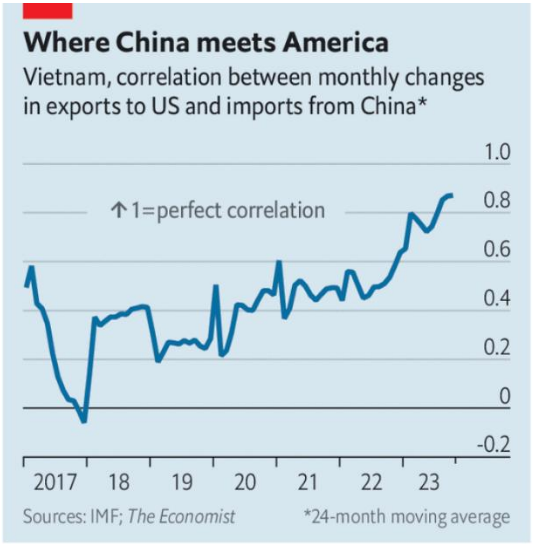
How Trump and Biden have failed to cut ties with China

It is hard to overcome economic incentives

Source: Economist (Feb 2024)

- More resilient or just more expensive, opaque, and complex?
- Lengthened but not more dense supply chain
- China's supply chain less visible but still extremely important





Why the Impact of Tariff is Mitigated?

Finance & economics | The levy paradox

Why Donald Trump’s tariffs are failing to break global trade

Six months on from “Liberation Day”, things look surprisingly rosy

Source: Economist
(Oct 2025)

- Tariffs are gentler than advertised with lags in implementation
- Carve-outs and exemptions (e.g., electronics and medicines)
- Large gap between statutory and actual levels
- Countries are diversifying their trade partners

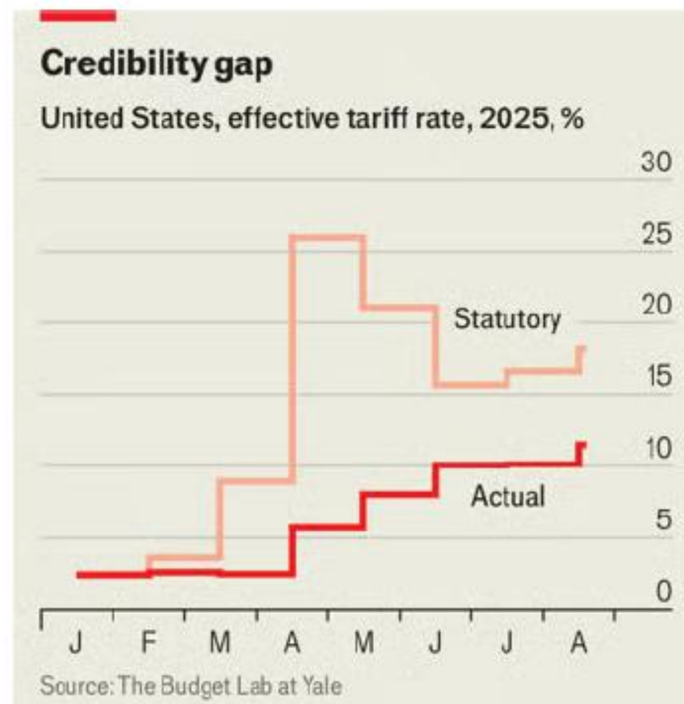


CHART: THE ECONOMIST

Future of Supply Chain Management

- With the development of technology, supply chain management is experiencing drastic changes with new challenges and chances

- (Extremely) high level of automation: e.g., “black” factory for electric cars; see the WSJ video at Link

- AI in manufacturing and supply chain: potentials in quality management, decision-making, cost reduction, and ESG
  - See the overview report at Link

- Large Language Models (LLM): smarter and faster decision-making, optimize logistics and adapt to market shifts
  - See a brief blog at Link

Summary

- Supply chain brings a new perspective in today's business world: think globally!
- Challenges in supply chain: uncertainty, swift changes, complex network
- Three A's wining strategies: Agility, Adaptability, Alignment
- Contract, double marginalization, and bullwhip effect
- Developments in supply chain: COVID-19 pandemic, US-China trade/tariff war, risk mitigation, automation and AI

ISOM 2700: Operations Management

Session 21. Coordinating Supply Chain: Risk-sharing Contracts

Yiwen Shen
Dept. of ISOM, HKUST
Fall 2025

Agenda

- **Nonalignment in supply chain: Wholesale price contract**

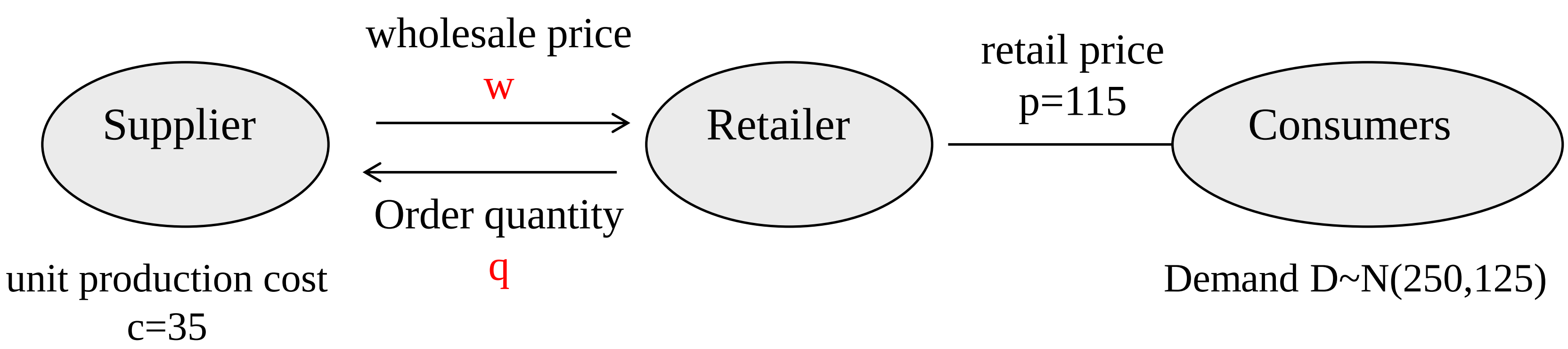
- Coordination with risk-sharing: Revenue sharing contract

Wholesale Price Contract: UV Sunglass

- Zamatia (supplier) makes sunglass at a cost of $35 and sells them to UV for $75
- UV (retailer) sells the sunglasses to customers for $115 and salvages leftover inventory for $25 per unit
- Customer demand is normal with mean 250 and standard deviation 125

Wholesale Price Contract

- Supplier (Zamatia) makes sunglasses at per unit cost $35, and sells it to retailer (UV) at the wholesale price w
- The retailer decides how many units of products to order, q, before the sales season, and pays the supplier the amount q*w
- The uncertain demand D~N(250,125) is realized
- The retailer gets $115 (retail price) for each unit sold and $25 (salvage value) for each unit of inventory



Supply Chain Optimum

- We first check the optimal order quantity for the entire supply chain
  - View retailer + supplier as a whole and ignore internal transfer
  - Similar to joint pricing learned before
  - It does not depend on what contract is used

- Supply-chain optimum: choose the ordering quantity to maximize the total expected profit of retailer and supplier

- What matters for the supply chain?
  - The production cost ($35), retail price ($115), and salvage value ($25)
  - Customer demand: normal distribution with mean 250 and standard deviation of 125

Supply Chain Optimum

- We solve the supply chain optimum by a newsvendor model
- Overstocking cost C_o = production cost– salvage value
  = 35 – 25 = 10
- Understocking cost C_u = retail price – production cost
  = 115 – 35 = 80
- Critical fractile = 80 / 90 = 0.89; z-score = 1.23
- Optimal order quantity = 250 + 125*1.23 = 404
- In order to maximize the total profit from customer, the optimal order quantity is 404

Optimal Ordering Quantity for Retailer

- Now, we focus on the optimal ordering quantity from the retailer under the wholesale contract
  - The retailer faces the wholesale price $75 from the supplier

- Retailer optimum: choose the ordering quantity to maximize its own expected profit

- Overstocking cost C_o = wholesale price – salvage value

  = 75 – 25 = 50

- Understocking cost C_u = retail price – wholesale price

  = 115 – 75 = 40

- Critical fractile = 40 / 90 = 0.44; z-score = -0.15

- Optimal order quantity = 250-0.15*125 = 232

Retailer Optimum vs. Supply Chain Optimum

- Under wholesale price contract, the order quantity is determined by the retailer
- However, the optimal order quantity for retailer (232) is different from (smaller than) the supply chain optimum (404)
- Misalignment of interest: the optimal decision for retailer may hurt the supply chain
- An example of double marginalization: it shows the challenge in supply chain coordination

Order Quantity: Critical Fractile

- In the newsvendor model, the ordering quantity is given by

  $$Q = \text{mean} + \text{standard deviation} \times \text{z-score}$$

- Given the mean and standard deviation of the demand, the ordering quantity is solely determined by the z-score

- The z-score is further obtained by the critical fractile (via the distribution table): $\Pr(Z \leq z) = Cu/(Cu+Co)$

- Thus, the critical fractile determines the ordering quantity in the supply chain

Limitation of Wholesale Price Contract

- Is it possible to achieve supply chain optimum by setting another wholesale price w?

- Denote new wholesale price by w> \$35 (otherwise wholesaler will lose money)
  - For retailer (UV): Cu = 115 - w, Co = w – 25
  - Critical ratio = $\dfrac{115 - w}{115 - w + w - 25} = \dfrac{115 - w}{90}$

- It is always smaller than the supply chain optimum as long as w is higher than the production cost \$35

$$\frac{115 - w}{90} < \frac{80}{90} \text{ given } w > \$35$$

Wholesale Price Contract: Findings

- Wholesale price contract always leads to lower than optimal order quantity for the supply chain
  - This happens for all feasible wholesale price levels

- Setting the wholesale price closer to the production cost can reduce the deviation from supply chain optimum
  - However, this hurts the profit of supplier

- This reveals the fundamental limitation of the traditional wholesale price contract

Root Cause: Lack of Alignment

- Why cannot we achieve optimum under wholesale price contract?

- Decisions are based on retailer's own margin, not the supply chain's margin (lack of alignment!)
  - Retailer: wholesale price, selling price, salvage value
  - Suppl chain: production cost, selling price, salvage value

- The supplier does not bear the uncertainty in customer demand (conflicting incentive!)

- The trade-off (order too many vs too few) for the retailer deviates from that for the supply chain

Expected Profit: Wholesale Contract

Back to the Zamatia and UV sunglass example, consider a wholesale price contract with wholesale price set at $75.

What is the expected profit of supplier and retailer respectively?

- For retailer, the profit for each sale is Cu = 115 – 75 = 40; the loss for each leftover is Co = 75-25 = 50

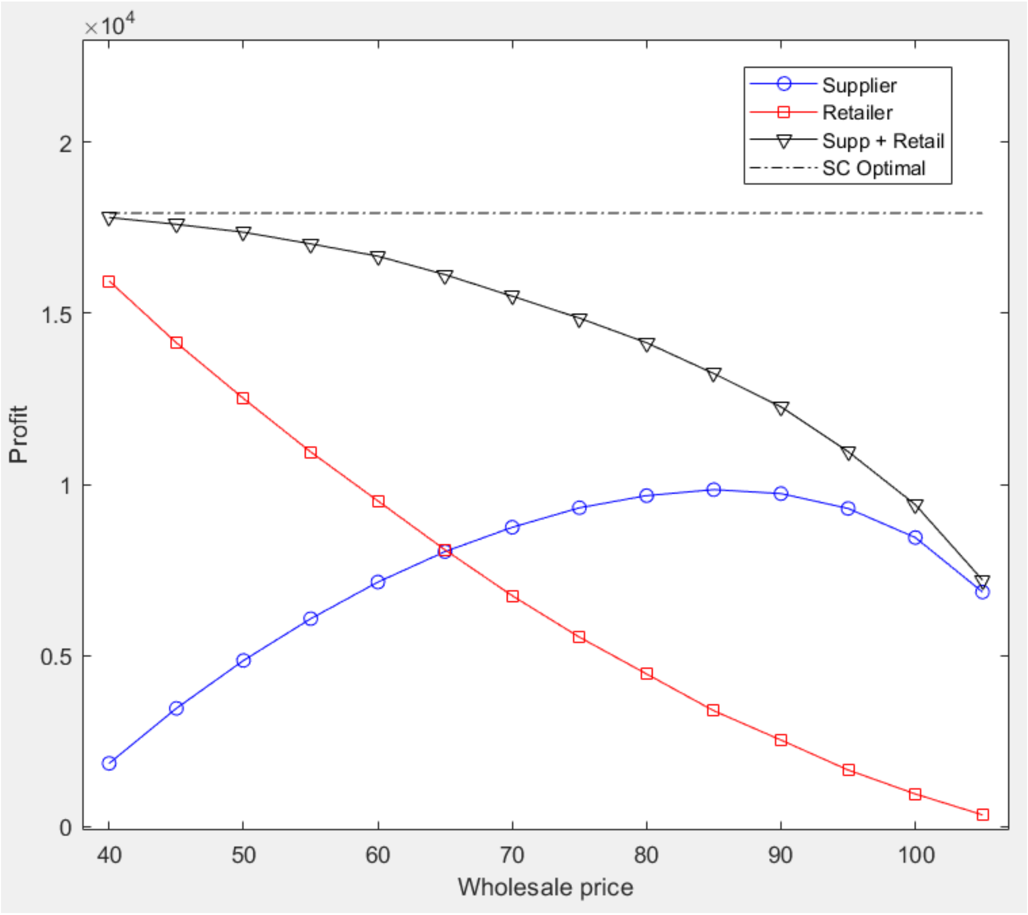
- Critical fractile = 40/90 = 0.44; from the distribution table we get z-score = - 0.13; the optimal order quantity = -0.15*125+250 = 232

- By inventory table, the expected leftover = 125*I(-0.15) = 125*0.3284 = 41
- Expected sales = 232 – 41 = 191

- Expected profit for retailer = 191*40 – 50*41 = 5590

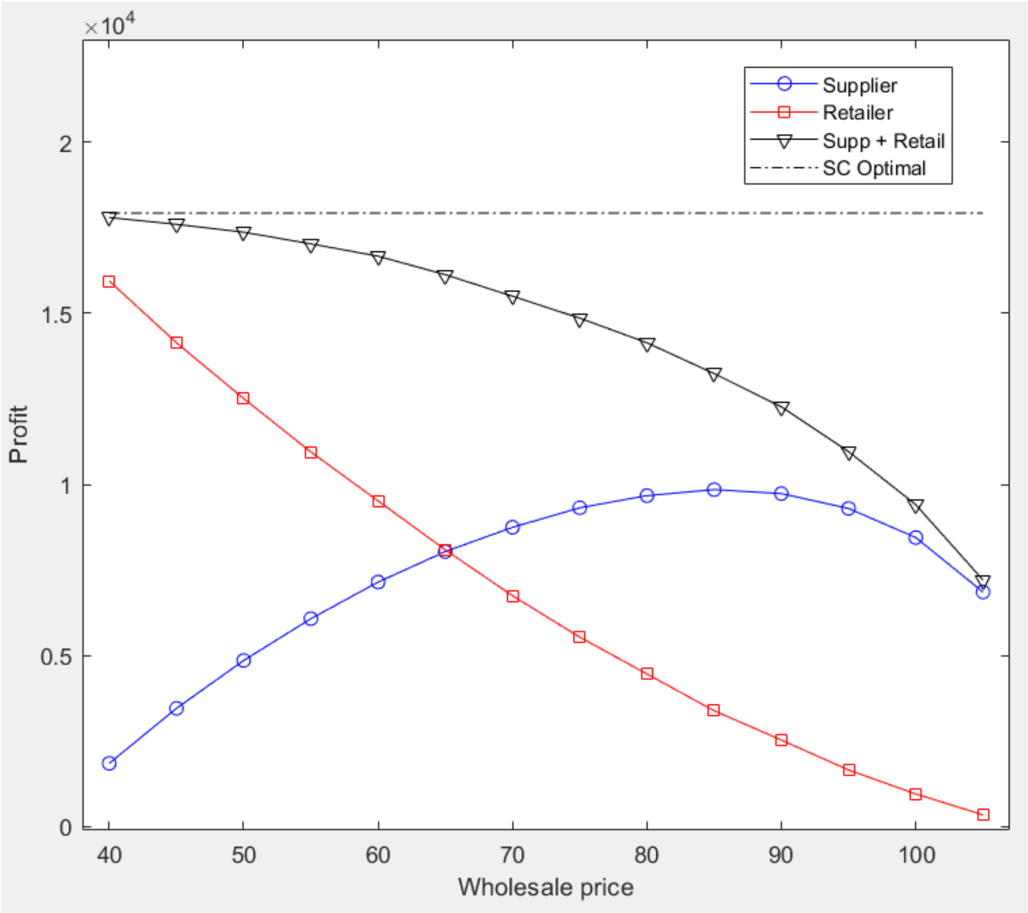
- Profit for supplier = 232*(75-35) = 9280

Distribution of Expected Profit



1. Vary wholesale price (x-axis)
2. Solve the optimal order quantity for retailer
3. Calculate the expected profit for retailer, supplier, and supply chain as a whole

Distribution of Expected Profit



At wholesale price x = $75

Profit for retailer = $5540
Profit for supplier = $9320

Total profit for supply chain = $14860

At supply chain optimum, total profit = $17930

Question: if you are the supplier, how would you set the wholesale price?

Agenda

- Nonalignment in supply chain: Wholesale price contract

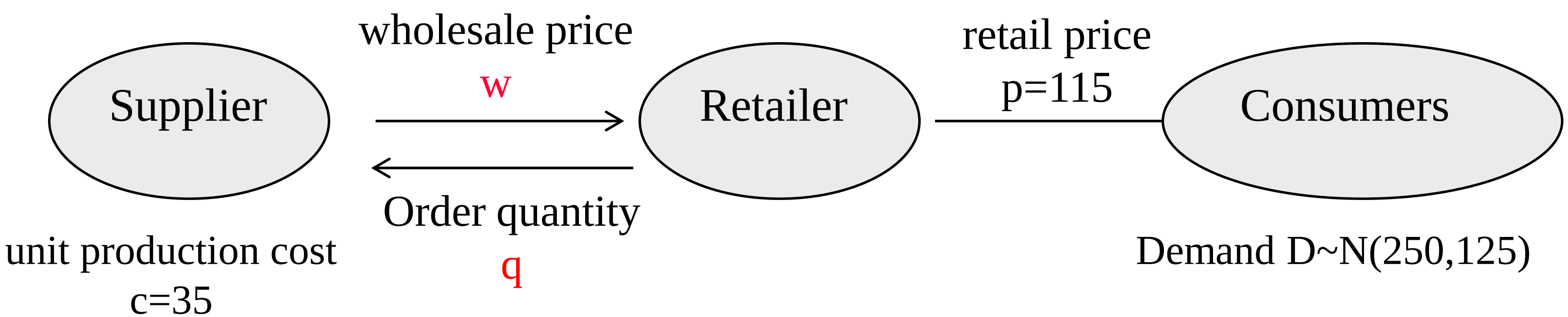
- **Coordination with risk-sharing: Revenue sharing contract**

Strategy: Contract Design

- We want to better design the contract to improve the efficiency of the supply chain
  - Hopefully achieve the supply chain optimum
  - A bigger pie may benefit everyone in the supply chain
- Remember the retailer still decides the order quantity based on its own benefit
  - i.e., the supply chain is still decentralized
- Key idea: share the risk or profit from customer demand with the supplier
  - Revenue sharing contract

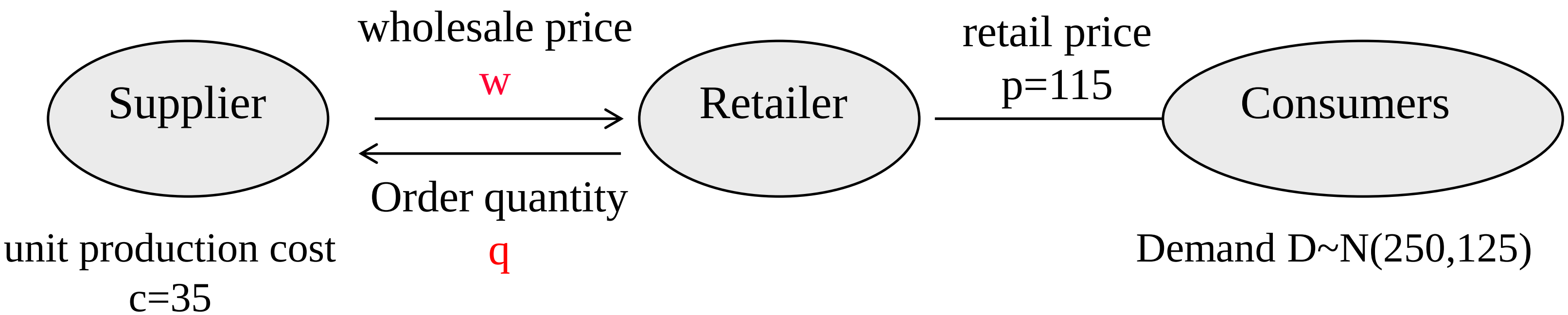
Revenue Sharing Contract

- The supplier charges a wholesale price w for each unit delivered to the retailer (assume w is higher than salvage value)
- For each unit of sales, the retailer pays the supplier a fixed percentage share y of the retail price
- The retailer decides how many units of products to order before the sales season



For every unit of sales, retailer gets 115(1-y) and supplier gets 115y

Revenue Sharing Contract



For every unit of sales, retailer gets 115(1-y) and supplier gets 115y

- Key idea: the revenue from sales is now shared between retailer and supplier
  - The ratio y controls how the revenue is shared
  - The retailer takes (w,y) as given parameters when ordering
  - When y = 0, we return to the wholesale price contract

Revenue Sharing Contract: Solution

- We still solve the order quantity by newsvendor model
- Given the wholesale price w and revenue share ratio, the retailer gets 115(1-y) for each unit sold
  - overstocking cost: $Co = \mathrm{w} - 25$
  - understocking cost: $Cu = 115(1 - \mathrm{y}) - \mathrm{w}$
- Critical fractile:

$$\frac{Cu}{Cu + Co} = \frac{115(1 - \mathrm{y}) - \mathrm{w}}{115(1 - \mathrm{y}) - 25}$$

- This determines the order quantity of the retailer

Optimal Price and Share Ratio

- What is the optimal level of price w and sharing ratio y that maximize the total supply chain profit?
  - i.e., achieving the supply chain optimum

- Note that the sharing between supplier and retailer does not affect total profit for supply chain (internal transfer)

- What matters is the order quantity from the retailer, which is determined by the critical fractile

- As long as the retailer orders the target quantity, we achieve the supply chain optimum

Optimal Price and Share Ratio

- Recall for the supply chain as a whole, the target critical fractile is 80/90 = 0.89, which leads to optimal order quantity
- The optimal (w, y) should make the retailer's critical fractile (thus order quantity) coincide with the supply chain's level
- This leads to

$$\frac{Cu}{Cu + Co} = \frac{115(1 - y) - w}{115(1 - y) - 25} = \frac{80}{90}$$

Optimal Price and Share Ratio

- For the supply chain to achieve its optimum, we need

$$\frac{Cu}{Cu+Co} = \frac{115(1-\mathrm{y})-w}{115(1-\mathrm{y})-25} = \frac{80}{90}$$

- After some simplification, we solve the optimal y for given w as

$$y = \frac{10\times 115 + 25\times 80 - 90w}{10\times 115}$$

- E.g., if wholesale price $w = 29$, we can solve $y = 46.9\%$

- If we use other wholesale price w, we can solve y accordingly to achieve the supply chain optimum

Structure of Revenue Sharing Contract

- In the revenue sharing contract, we can choose the wholesale price and sharing ratio y to achieve the supply chain optimum
- For y to be feasible (in [0,1]), we need the wholesale price w to stay between the salvage value and the production cost
- That is, the supplier first loses money by selling to the retailer at first, but gets compensated from revenue sharing
- This structure of optimal revenue sharing contract can be proved (the proof is not required)

Achieving Supply Chain Optimum

- With proper (w,y), the total supply chain profit (retailer + supplier) can be maximized by the retailer's ordering quantity

- Revenue sharing contract allows the decentralized supply chain to achieve the same profit as the centralized supply chain
  - Eliminate double marginalization

- Economic intuitions:
  - The revenue sharing contract motivates the retailer to order the "right" quantity for the entire supply chain
  - The supplier now bear the profit and uncertainty in customer's demand
  - The supplier and retailer's incentives are better aligned

Expected Profit: Revenue Sharing Contract

In the UV example, suppose we use a revenue sharing contract with wholesale price w = 29 and sharing ratio y = 47%.
What is the expected profit of retailer and supplier?

- For retailer, the profit for each sale is Cu = 115*(1-y) – w = 32; the loss for each leftover is Co = 29-25 = 4

- Critical fractile = 32/36 = 0.89; from the distribution table we get z-score = -1.23; the optimal order quantity = 1.23*125+250 = 404

- By inventory table, the expected leftover = 125*I(1.23) = 125*1.2827 = 160
- Expected sales = 404 – 160 = 244

- Expected profit for retailer = 32*244 – 4*160 = 7168

- Expected profit for supplier = 404*(29-35) + 244*115*0.47 = 10764
- The first part is from the wholesale stage, the second part is from the revenue sharing stage

Profit with Revenue Sharing Contract

- Set wholesale price $w = 29$, and the corresponding revenue sharing ratio $y = 47\%$
- We can compute the profit for supplier and retailer

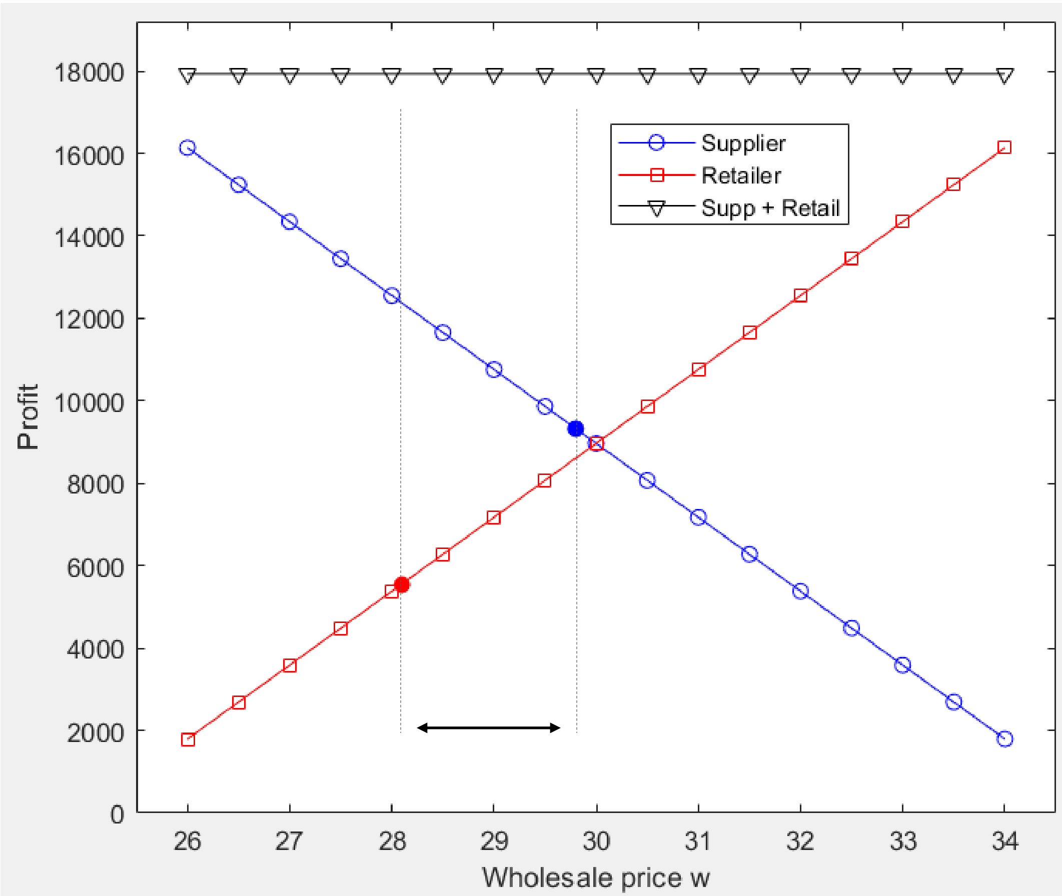
	Wholesale Price (x = $75)	**Revenue Sharing (w = $29, y = 47%)**
Supplier's Profit	9320	10764
Retailer's Profit	5540	7168
Supply chain's profit	14860	17930 (optimum)

- Both profits are higher than the wholesale price contract case: win-win situation

Distribution of Profit

- For a given wholesale price w, we can set a corresponding sharing ratio y to achieve supply chain optimum
- All such (w,y) maximize the total profit of supply chain (retailer + supplier) by definition
- However, changing (w,y) leads to a different distribution of the total profit between retailer and supplier
- Both total profit and distribution of profit matter in supply chain management

Distribution of Profit



- Plot the retailer and supplier's profit at different price w
- The value y is set accordingly to get optimal order quantity
- The value of w determines the distribution of profits between supplier and retailer
- There is a range of w (and corresponding y) such that both retailer and supplier are better off
- (the solid points are the level from wholesale contract)

Distribution of Profit

- Under the revenue sharing contracts that achieve supply chain optimum, it is possible that
  - the expected profit of one player becomes lower than before
  - then, the player is not willing to switch to the revenue sharing contract

- There always exists a specific range of wholesale price w (and corresponding y) such that
  - supply chain optimum is achieved
  - and simultaneously both players get expected higher profit
  - a win-win situation for the system and both players

Summary of Supply Chain Contracts

	Supply chain optimum	**Wholesale contract**	**Revenue sharing contract**
Who places order	Central planner	Retailer	Retailer
Contract parameters	N/A	Wholesale price	Wholesale price, sharing ratio
Overstocking cost	Production cost – salvage value	Wholesale price – salvage value	Wholesale price – salvage value
Understocking cost	Retail price – production cost	Retail price – wholesale price	Revenue for retailer – wholesale price
Achieve SC optimum	N/A	Impossible	Possible

Formula for Revenue Sharing Contract

- Suppose the production cost is p_{prod}, selling price is p_{sell}, and salvage value is p_{salv}. They are given by the problem set-up.

- For the supply chain optimum to be achieved, the wholesale price w and share ratio y should satisfy

$$\frac{p_{sell}(1-y)-w}{p_{sell}(1-y)-p_{salv}}=\frac{p_{sell}-p_{prod}}{p_{sell}-p_{salv}}$$

- By above we can solve the optimal wholesale price w and share ratio y satisfy

$$y=\frac{p_{prod}(p_{sell}-p_{salv})\ -(p_{sell}-p_{salv})w}{p_{sell}\left(p_{prod}-p_{salv}\right)}$$

- Note that we need $p_{salv} < w < p_{prod}$ for y to be between 0 and 1

Expected Profits in Supply Chain Contracts

- Wholesale price contract:
  - Retailer: use newsvendor model formula
  - Supplier: order quantity × (wholesale price – production cost)

- Revenue sharing contract:
  - Retailer: use newsvendor model formula with new overstocking and understocking cost
  - Supplier: order quantity ×(wholesale price – production cost) + selling price × sharing ratio for supplier × expected sales

Other Types of Risk-sharing Contracts

- (Not required)
- Return contract: the retailer can sell back the leftovers to the supplier at a price higher than the salvage value
- Quantity contract: the retailer receives a discount on all units if the ordering quantity exceeds a thresholds (encourage buying)
- Option contract: the retailer pays the supplier an upfront amount to build capacity and buys an option to purchase later
  - Motivate capacity building before selling season
- Quantity flexibility contract: the buyer/supplier agree to buy/sell the quantity within a certain range of forecast (e.g., 75% to 125%)
  - Capacity planning under demand uncertainty
- Price protection contract: the supplier compensates the distributor for any price reduction on remaining inventory
  - Used in the tech sector to mitigate the price drop due to product updates

Knowledge Points

- Concept of double marginalization

- Supply chain optimum
  - Definition and interpretation
  - Solve optimal order quantity and performance metrics

- Wholesale contract:
  - Solve optimal order quantity
  - Calculate expected profit and performance metrics
  - Compare it with supply chain optimum

- Revenue-sharing contract:
  - Solve optimal order quantity under a revenue-sharing contract
  - Find the optimal sharing ratio to achieve supply chain optimum

Takeaway

- Two aspects in designing supply chain

- Total profit pie:
  - What behavior maximizes system-wide efficiency?
  - How can contract provide incentives for this behavior?
  - Who is at the best position to absorb risks?

- Distribution of the profit pie:
  - Contract terms must be set so that they are attractive to both firms
  - The distribution depends on the bargain power of the participants

ISOM 2700: Operations Management

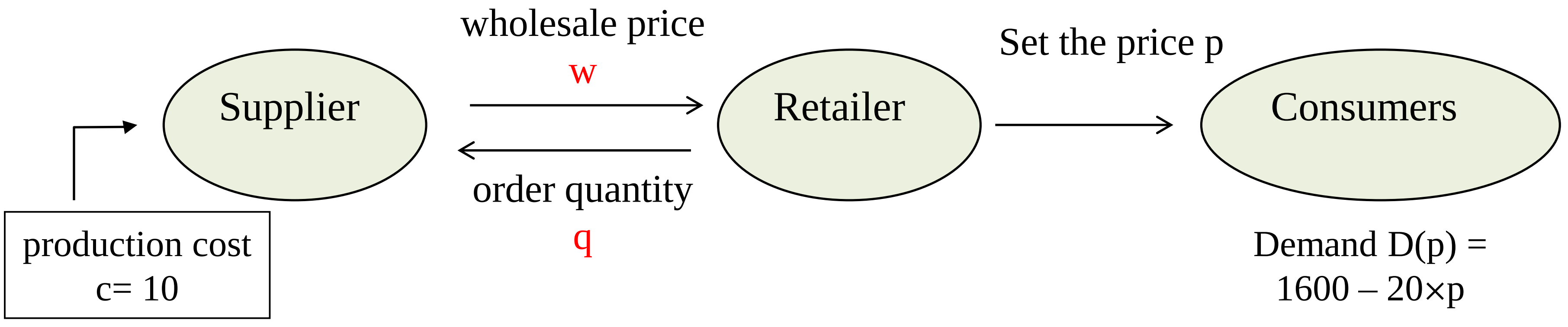
Session 22. Price-based RM in Supply Chain & Behavioral OM

Yiwen Shen
Dept. of ISOM, HKUST
Fall 2025

Price-based RM in Supply Chain

Consider a supply chain with a supplier and a retailer. The retailer buys the products from the supplier at the wholesale price and then sells the product to customers at the retail price. The retail price and order quantity are determined by the retailer. The wholesale price is set by the supplier.

For a retail price p, the demand is given by $D(p) = 1600 - 20 \times p$. The production cost is 10 HKD per unit.



Supply Chain Optimum

Q: If we view the supply chain as a whole, what is the optimal retail price and the total profit?

Now you only need to consider the customer demand and production cost

Profit of the supply chain = $(1600 - 20p) \times (p - 10)$

$= -20p^2 + 1800p - 16000$

To maximize it, optimal $p = 1800/(2\times 20) = 45$

Demand = $1600 - 20\times 45 = 700$

Total profit = $700\times 35 = 24500$

Price-based RM in Supply Chain

Q: Suppose the wholesale price is set at 30, what is the optimal retail price? What is the profit of the retailer and supplier?

Profit of retailer = $(1600 - 20p) \times (p - 30) = -20p^2 + 2200p - 48000$

Optimal price $p = 2200/(2\times20) = 55$

Order quantity $= 1600 - 20\times55 = 500$

Profit of retailer $= 500\times(55-30) = 12500$

Profit of supplier $= 500\times(30 - 10) = 10000$

Total profit = 12500 + 10000 = 22500

Optimal Response to a Wholesale Price Level

Q: Suppose the wholesale price is set at w what is the optimal retail price? What is the profit of the retailer and supplier?

Profit of retailer $= (1600 - 20p)\times(p - w)$

$= -20p^2 + (1600+20w)p - 1600w$

Optimal price $p = (1600+20w)/(2\times20) = 40 + w/2$

Order quantity $= 1600 - 20\times(40+w/2) = 800 - 10\times w$

Profit of retailer $= (800 - 10\times w)\times(40+w/2-w)$

$= 5w^2 - 800w + 32000$

Profit of supplier $= (800 - 10w)\times(w - 10)$

$= -10w^2 + 900w - 8000$

Optimal Wholesale Price for Supplier

Q: If you are the supplier, what wholesale price you will set? What is the corresponding total profit?

Profit of supplier = $(800 - 10w) \times (w - 10)$

$= -10w^2 + 900w - 8000$

To maximize it, optimal $w = 900/(2\times10) = 45$

Retail price $p = 40 + w/2 = 62.5$

Order quantity $= 800 - 10\times w = 350$

Profit of retailer = 6125

Profit of supplier = 12250

Total profit = 12250 + 6125 = 18375

Limitation of Wholesale Contract

Q: Can we achieve supply chain optimum using the wholesale price contract?

Given wholesale price w, the retail price is p = 40+w/2

For supplier to make money, the wholesale price must be higher than 10, i.e., w>10

Then, we always have p = 40+w/2 > 45

Thus, the supply chain optimum cannot be achieved. The retail price will be higher than that

This again reflects the negative impact of double marginalization

Revenue Sharing Contract

Q: Suppose the supplier and retailer now enter a revenue sharing contract. The wholesale price is w and the retailer share a proportion y of the revenue to the supplier.

To achieve the supply chain optimum, what should be the w and y?

To achieve the supply chain optimum, the retailer needs to set the retail price at p = 45, which maximizes the total profit

Revenue Sharing Contract

Q: Suppose the supplier and retailer now enter a revenue sharing contract. The wholesale price is w and the retailer share a proportion y of the revenue to the supplier. To achieve the supply chain optimum, what should be the w and y?

$$\text{Retailer profit} = (1600 - 20p) \times (p\text{-}w) \times (1\text{-}y)$$
$$= [-20p^2 + (1600+20w)p - 1600 \times w] \times (1\text{-}y)$$

Note that the coefficient (1-y) does not affect the optimal retail price, thus we have optimal $p = 40+w/2$

To achieve supply chain optimum, we need retail price $p = 45$, thus the wholesale price should be set at $w=10$, i.e., the production cost

We do not have limitation on y, it controls how the profit is split between supplier and retailer

Distribution of Profit

Q: Comparing with the original wholesale price contract with w=30, what is the range of sharing ratio y such that both retailer and supplier are willing to switch to the revenue sharing contract?

In the original wholesale price contract with w=30, the retailer profit = 12500; the supplier profit = 10000

In the revenue sharing contract with w=10, the total profit equals 24500 (supply chain optimum)

Retailer profit = $24500\times(1-y)$; supplier profit = $24500\times y$

For retailer to switch: $24500\times(1-y) > 12500 \Rightarrow y < 49\%$

For supplier to switch: $24500\times y>10000 \Rightarrow y> 40.8\%$

Thus, we require $40.8\% < y < 49\%$

Summary

- Price-based RM problem in a supply chain with deterministic demand
- The final demand is a linear function of the retail price; there is a production cost for supplier
- Q: In traditional wholesale contract, what should be the optimal retail price for a given wholesale price?
- Q: In traditional wholesale contract, how should the supplier set its optimal wholesale price to maximize its own profit?
  - Need to consider the response of the retailer (order quantity)
- Q: what should be the optimal retail price to maximize the total supply chain profit (supply chain optimum) ?

Summary

- The traditional wholesale price contract cannot achieve supply chain optimum --- the wholesale price will be set at a high level
- Supply chain optimum can be achieved by a revenue sharing contract
- In the revenue sharing contract, the wholesale price is set at the production cost, the sharing ratio controls the profit distribution
- There is a range of sharing ratio such that both supplier and retailer are better off
  - Both have higher profit in the revenue sharing contract

Formula

- Suppose the demand for product is given by D(p) = a-bp, where p is the retail price. The production cost is c per unit.

- In the traditional wholesale contract, for a given wholesale price w, the optimal retail price is given by

$$p^* = \frac{1}{2b}(a + bw)$$

- For the supplier, its optimal wholesale price (after accounting for the retailer's response) to maximize its profit is given by

$$w^* = \frac{1}{2b}(a + bc)$$

- For the supply chain optimum, the retail price should be

$$p^* = \frac{1}{2b}(a + bc)$$

Formula

- For the revenue sharing contract to achieve supply chain optimum:

- The wholesale price should be set at w=c (production cost)

- The retail price should be set at the supply chain optimum level

$$p^* = \frac{1}{2b}(a + bc)$$

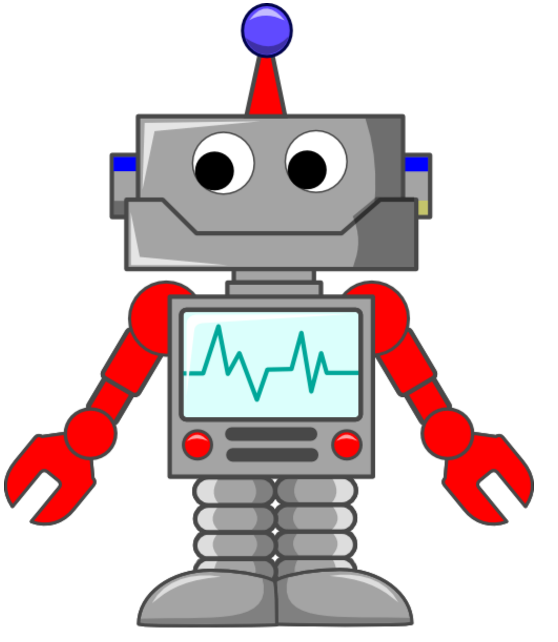
- The sharing ratio y controls the distribution of profit

Introduction to Behavioral OM

- **Intro to Behavioral OM**
- Behavioral OM: Judgement Under Uncertainty
- Behavioral OM: Risky Choice and Preference
- **(This part is NOT required in the exam)**

Traditional Economics

- Assumes humans are “optimal” decision making machines:
  - Reasoning capacity is infinite
  - People act in their long-term interest
  - Markets are perfectly efficient
  - Every act is entirely selfish
  - On average, beliefs are unbiased
  - Incentives always matter
  - Over time, people get it right



Traditional OM

- Assumes humans in a process act like perfect machines
  - Workers have a well defined and fixed "processing speed"
  - Customers arrive randomly, have rational rules for waiting/buying
  - Process and job design matter more than "fairness"

Behavioral Econ & OM

- Making fewer unrealistic assumptions about human decision makers

- The standard approaches are powerful, but are not the only valid approach

Big Ideas of Behavioral Research

- People use simple decision heuristics based on limited information to make most decisions
- People take the easy path, succumb to temptation, and follow habits
- Social forces like fairness and status are major motivations
- Strategic sophistication is limited

Outline

- Intro to Behavioral OM
- **Behavioral OM: Judgement Under Uncertainty**
- Behavioral OM: Risky Choice and Preference

Judgment Biases

Availability	Representativeness	Confirmation
• Vividness/recency • Retrievability (memory or search) • Inattentional blindness	• Base rate neglect • Sample size Insensitivity • Gambler's fallacy/LoSN • Reversion to the mean • Conjunction fallacy	• Anchoring bias • Conjunctive/disjunctive events bias • Over-precision • Hindsight bias

Motivating Examples

Given the following description of Tim (in HK), what job do you think he is more likely to do?

"Tim is a detail-oriented and careful person, he is very responsible for his work, and he arranges everything in an orderly manner"

A. Doctor
B. Investment banker
C. Business professor
D. Staff in a firm

Availability Bias

Availability bias:

- Assess the frequency, probability, causes of an event by the information readily "available" in memory

- Rely disproportionately upon the most readily available data.

Availability Bias: Examples

- Vivid and recent information:
  - Think these events are more frequent/numerous
  - Biases quantity and probability estimates

- Example: people are more likely to purchase earthquake insurance immediately after an earthquake

Availability Bias: Examples

- Retrievability:
  - Easier to retrieve the information from memory/search strategy and use this more available information
  - Biases quantity and probability estimates
- Example: Hiring managers tend to search candidates similar to their background first

- Inattentional Blindness:
  - Use information related to what you are focusing on and do not use not others
  - Biases quantity and probability estimates about others
- Example: a policeman chasing a suspect may fail to notice a fight along the road

Representativeness Bias

Representativeness bias: making judgments based on the degree to which an event is

- similar in essential characteristics to its parent population
- reflects the salient features of the process by which it is generated

"Things should look like what they represent"

Representativeness Bias

- Base rate neglect:
  - Assume signals will tend to look like what they are signaling
  - Ignore base rate relative to the (more vivid) signal information
- Example: people may overestimate the likelihood of a rare disease given positive test result

- Reversion to the mean:
  - Think this period's signal is representative of the underlying true expected value
  - Fail to predict reversion to the mean in next periods
- Example: overestimate the distribution of demand if the current period demand is high

Representativeness Bias

- An example for base rate neglect bias
- Suppose now it is estimated that 1% of population has infected Covid. Steve took the PCR test and found positive.
- The PCR test is not fully accurate:
    - if you are truly positive: 90% of chance to test positive, 10% chance to test negative
    - if you are truly negative: 95% of chance to test negative and 5% chance to test positive
- In your mind, what is the probability that Steve is truly positive?
- A: >= 80%, B: 50% to 80%, C: 20% to 50%, D: 10% to 20%

Confirmation Bias

Confirmation bias:

- search for, interpret, and favor information in a way that confirms or supports one's prior beliefs or values
- ignore contrary information or interpret ambiguous evidence as supporting their existing attitudes.

Confirmation Bias

- Anchoring Bias:
  - Selectively testing numbers near the anchor or search for information to justify that anchor
  - Stop testing when the number is plausible
- For example, to estimate the average price of new German car, people tend to refer to the price of influential Germany brands

- Over-precision:
  - Search memory for evidence that confirms your guess and don't search enough for disconfirming evidence
  - Overestimate accuracy of our knowledge and the truth of tentative hypothesis.
- For example, people don't put enough weight on others' advices

Outline

- Intro to Behavioral OM
- Behavioral OM: Judgement Under Uncertainty
- **Behavioral OM: Risky Choice and Preference**

Risky Choice

Consider a disease that is threatening the lives of 600 people

Program A: 200 people will be saved

Program B: there is a one-third probability that 600 people will be saved and a two-thirds probability that no people will be saved

Which of the two programs will you favour?

I. Program A
II. Program B

Risky Choice

Consider a disease that is threatening the lives of 600 people

Program C: 400 people will die

Program D: there is a one-third probability that no one will dies and a two-thirds probability that 600 people will die

Which of the two programs will you favour?

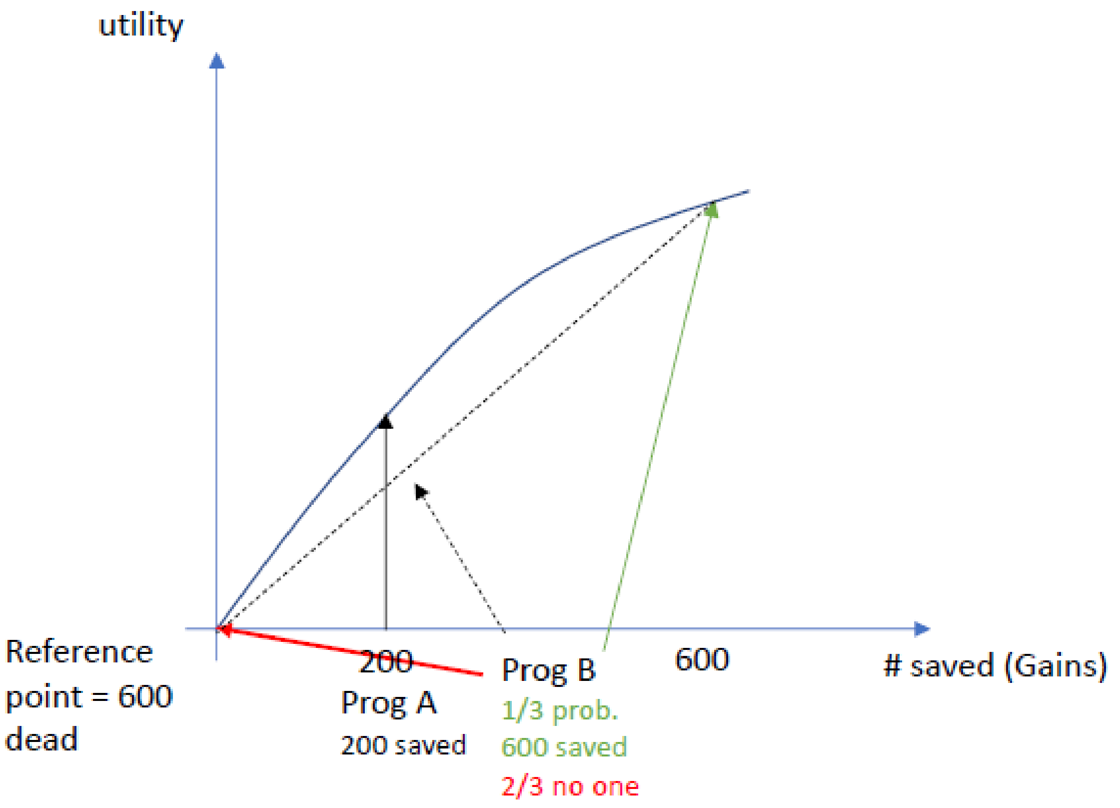
I. Program C
II. Program D

Observation

- The options A and C are the same, so are options B and D
- Usually people favour A over B, but D over C
- Why people make different choices?
  - Framing of the problem matters
  - gain vs. loss frame
- Observation: Risk-aversion in gains, risk-seeking in losses
- Questions: how to capture such behaviour by the utility function?

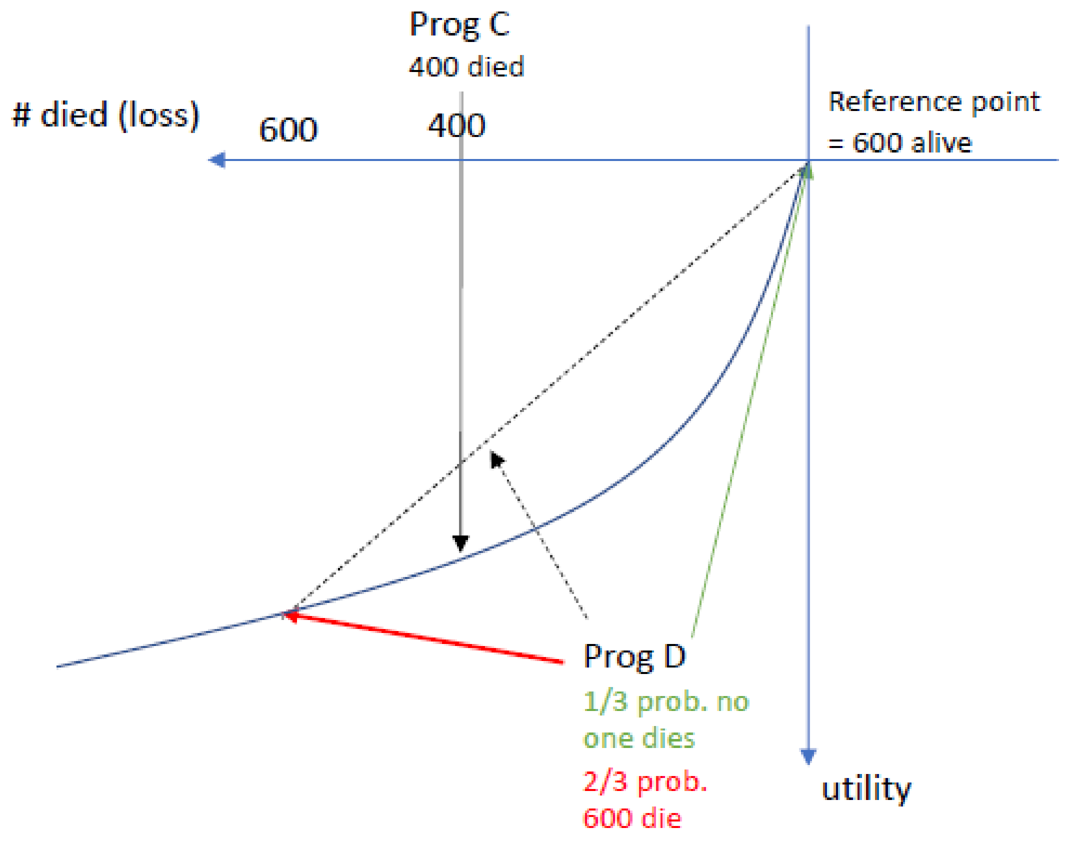
Utility Function: Gain

We use a concave function to capture the “risk-aversion” in gains



Utility Function: Loss

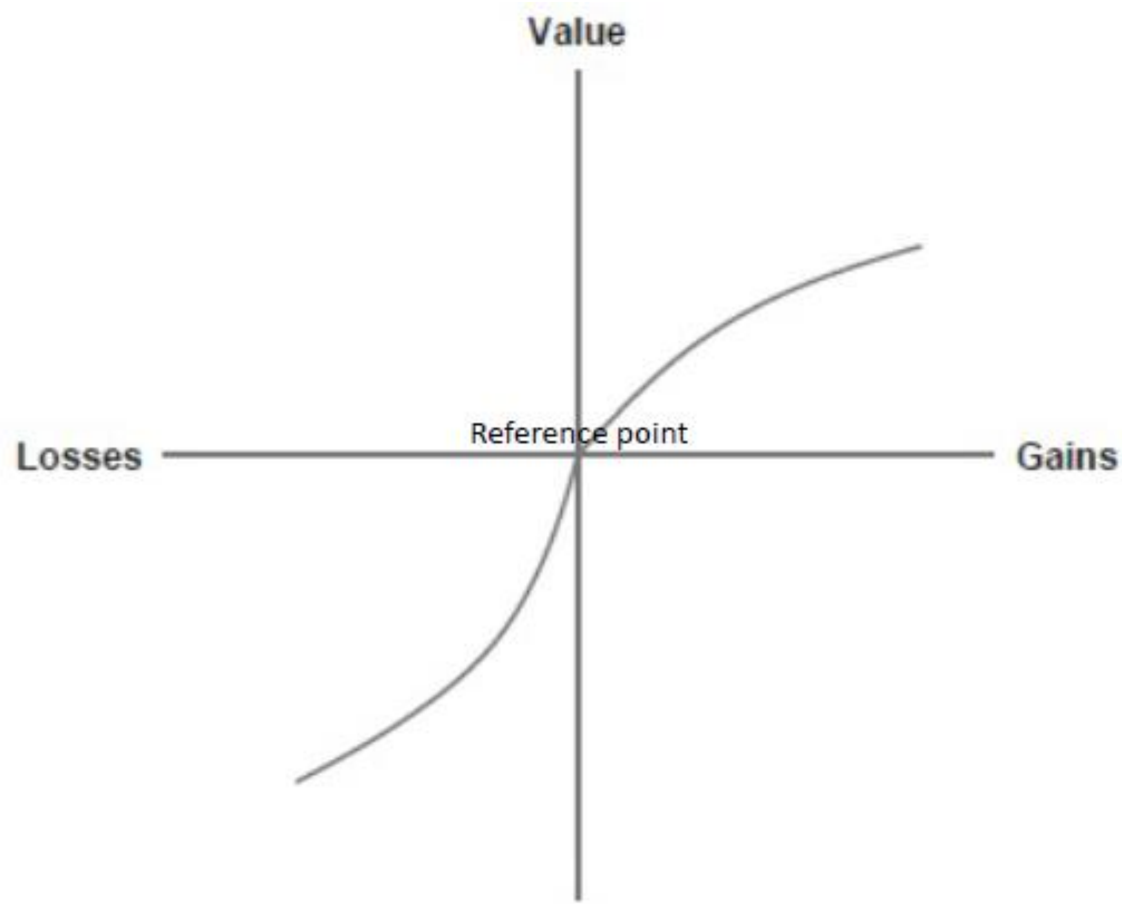
We use a convex function to capture the "risk-seeking" in loss



Utility Function Curvature

Combine the two parts, we get a utility function with curvature

The reference point matters, but it is not always obvious



Other Topics in Preference Study

- Probability weighting
- Utility function kink
- Mental accounting
- Time preferences