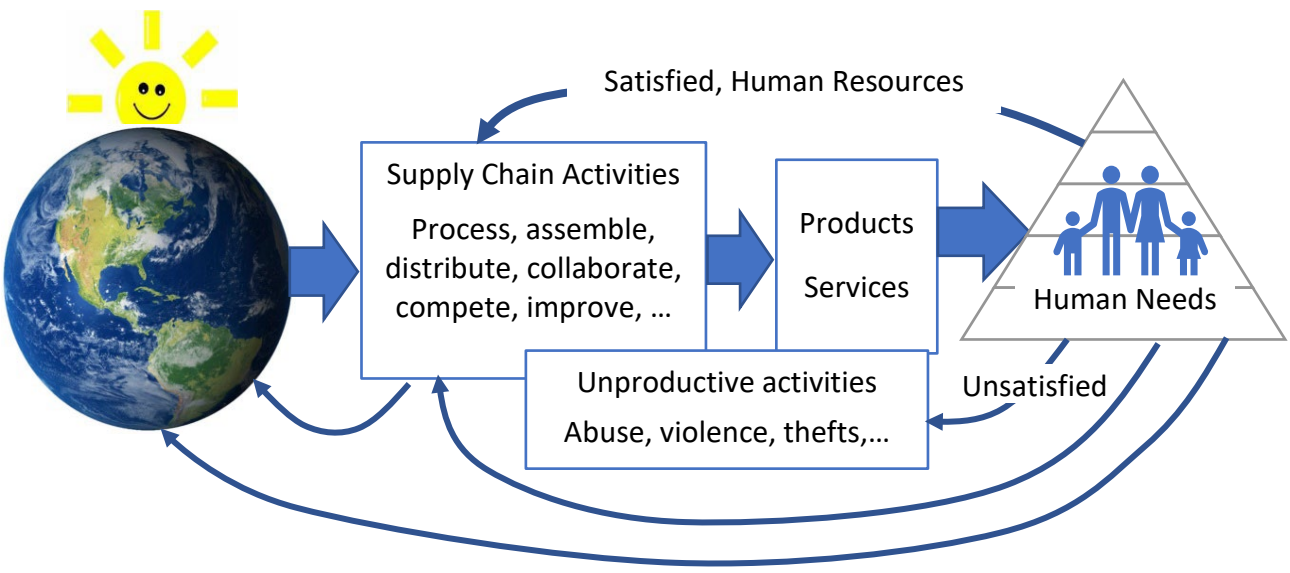
SUPPLY CHAIN WITH ECONOMICS AND HUMAN NEEDS

Design and Analysis of Supply Chain Operations and Strategies



Chen Zhou

This book is based on an engineering course in supply-chain economics. It draws on experience in teaching supply chain, operations, facility tours and projects from over 20 years. The context is supply chain with a focus on cost reduction and efficiency. It differs from other supply chain books by adopting a few economic models for collaboration, competition, pricing, scale, scope, information asymmetry, incentives and moral hazard.

A unique feature of this textbook is the link between supply chain and human needs. A low-cost, efficient, coordinated, and competitive supply chain serves human needs well—but not always. Human needs such as nutrition, health and livelihood, provide a concrete connection between the supply-chain strategies and their impacts on profit, society and environment. The externalities in economics can explain the misalignments in certain cases.

This textbook includes qualitative discussions on design and operation strategies and quantitative models for design and analysis. The quantitative models require a knowledge of algebra, basic calculus, and statistics. The practitioners who do not have some of the mathematical background can still follow the qualitative discussions and the results from the analysis.

Chen Zhou, PhD, has been a faculty member in the Stewart School of Industrial and Systems Engineering at Georgia Institute of Technology since 1989. His research and teaching interests are in supply chain, sustainability, logistics and warehousing.

BOOK REVIEW

In his book Supply Chain with Economics and Human Needs, Professor Chen Zhou takes a uniquely human centered approach to covering key topics in contemporary and Next Generation supply chain design and operations that include the economies of collaboration, scale, sharing, scope, and speed; the benefits of principal-agent modeling; competition and its link with game theory; pricing; logistics (transportation and inventory); uncertainty, risk, and risk mitigation; and innovative contracting. The textbook captures lessons learned about supply chains during the pandemic. The book is useful for the business school or engineering undergraduate with a good background in algebra, basic calculus, and statistics. An excellent textbook with a case study motivation to serve as an upper-level undergraduate course textbook or as a reference for both undergraduate and master's level courses in applied supply chain design and operations.

Chelsea White, PhD, Professor and Schneider National Chair in Transportation and Logistics, Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology.

ISBN 979-8-765-54561-4

Copyright © 2022 by Chen Zhou

ISBN 979-8-765-54561-4

V 1

Cover Design by Jenny S. Zhou

First paperback edition March 25, 2022

For more information, contact author at cz3@gatech.edu.

isye.gatech.edu/users/chen-zhou

SUPPLY CHAIN WITH ECONOMICS AND HUMAN NEEDS

DESIGN AND ANALYSIS OF SUPPLY CHAIN OPERATIONS AND STRATEGIES

By Chen Zhou
Stewart School of Industrial and Systems Engineering
Georgia Institute of Technology
Atlanta, GA 30332

April 2022

Table of Contents

1. INTRODUCTION to SUPPLY CHAIN for HUMAN NEEDS 1
    1. Overview of the Supply Chain, Economics, and Human Needs 1
    2. Supply Chain 4
        2.1 The production chain of a calculator 4
        2.2 Distribution of a calculator 4
        2.3 Supply chain defined by CSCMP 5
    3. Microeconomics in the Supply Chain 5
        3.1 Collaboration 6
        3.2 Competition 6
    4. Supply to Human Needs 7
    5. The Outcome of the Supply Chain to Human Needs 9
        5.1 Supply chain achievements 9
        5.2 Abundance and scarcity 10
        5.3 The assumption of rational choice 10
        5.4 Human needs and human resources 10
        5.5 Pricing signals and human needs 11
        5.6 Economic externalities 11
    6. Examples of Misalignment between Human Needs and Pricing 11
        6.1 Food deserts 11
        6.2 High cost and low performance healthcare supply chain in the United States 12
        6.3 Other examples of misaligned supply chains and human needs 17
    7. The Cost of Unsatisfied Human Needs 18
    8. The Change 19
        8.1 Human needs and triple bottom line in education 19
        8.2 Corporations 20
        8.3 Government 20
        8.4 Government or market 20
    9. Organization of the Book 20
        9.1 Structure of the chapters 20
        9.2 The content in the chapters 22

Acknowledgements 22
Review questions 22
2. COLLABORATION in SUPPLY CHAIN 23
1. Broad Perspectives in Collaboration and Competition 23
1.1 Collaboration challenges with the suppliers 24
1.2 Collaboration challenges with the clients 24
1.3 Coopetition 24
1.4 Factors that impact the collaborations 25
2. The Dry Pasta Supply Chain 26
2.1 Dry pasta production and distribution 26
2.2 Producer marketing 27
2.3 The performance indicators 27
2.4 The problems with high order variability 28
2.5 A proposal and its reception 28
2.6 Discussions 28
3. Methods to Improve Collaboration 28
3.1 Identify opportunities 29
3.2 Challenges in incentive alignment 29
3.3 Identify the differences 31
3.4 Possible solution strategies 31
3.5 Invest in information-sharing 33
4. Examples 33
4.1 Waste management in conference centers 33
4.2 Barilla's case study 35
4.3 Other examples in the HBR Article 36
4.4 Various forms of supply-chain contracts 37
5. Collaboration and Human Needs 37
3. PRINCIPAL-AGENT MODEL 39
1. Information and Behavior 39
1.1 Public information 39
1.2 Private information 39

1.3 Laws on privacy 40
1.4 The cost of information gathering, organizing, and sharing 40
1.5 Information asymmetry 41
1.6 Information sharing 42
1.7 Signaling 42
1.8 Adverse selection 42
1.9 Moral hazard 42
2. The Principal-Agent Model 43
2.1 Principal and agent 43
2.2 Important milestones in contracting process 43
2.3 Agent actions 44
2.4 Nature or non-controllable factors 44
2.5 Outcomes, pay-offs and objectives of the contract 45
3. Quantitative Principal Agent Model 45
3.1 The objective function for the principal 45
3.2 The constraints for the principal 46
3.3 The objective function for the agent 46
3.4 Constraint for the agents 47
4. Two-Level Principal-Agent Model 47
4.1 Objective functions and constraint for the principal 48
4.2 The objectives and constraints for the agent 48
4.3 An example of a risk-neutral model 49
4.4 Risk-averse agent 54
4.5 Agents with different levels of other opportunities 55
4.6 Effort observability 55
4.7 Two-level principal-agent model summary 56
5. Toward Vendor-Managed Inventory 56
5.1 The key characteristics of VMI 56
5.2 Some examples of VMI 58
5.3 Principal agent model and VMI 59
6. Principal Agent Model and Human Needs 59

6.1 Strategies of the conference center waste management project 59
6.2 Singapore's strategy to align finance and human's health needs 59
6.3 A U.S. example to align finance with human health needs 60
6.4 Issues in the U.S. healthcare system 61
4. COMPETITION in SUPPLY CHAIN 63
1. Porter's Other Three Forces in Competition 63
1.1 Threat of new entrants 63
1.2 Threat of substitutes 63
1.3 Rivalry among existing competitors 64
2. Monopoly Maximum Profit Supply Quantity 64
2.1 Reverse demand model 64
2.2 Monopoly maximum profit model 65
3. Cournot Competition in Duopoly 70
3.1 Modeling and solutions 70
3.2 Playing the Cournot game 71
3.3 Nash equilibrium 72
4. Cournot Competition and the Prisoner's Dilemma in Game Theory 75
4.1 Prisoner's dilemma 75
4.2 Game representation in strategic form 76
4.3 Cournot competition is an application of the prisoner's dilemma 77
4.4 The concept of dominance and dominated action 77
4.5 Finding Nash equilibrium based on elimination of dominated strategies 77
5. Cournot Competition Extensions 78
5.1 *n*-players 78
5.2 Duopoly with cost reduction 79
5.3 Importance of information 82
5.4 Tacit Collusion 82
6. Bertrand Competition 82
6.1 Modeling 82
6.2 Playing the Bertrand game 84
6.3 Nash equilibrium 85

6.4 Bertrand extension on different costs 85
7. Competition with Differentiation 85
7.1 Dimensions to differ 86
7.2 Cost and benefit to differentiate 86
7.3 The law of minimum differentiation in location 87
7.4 The Bertrand competition with differentiated price 88
8. Summary of Economies of Competition 91
8.1 Practical considerations 92
8.2. Strategies in a supply chain in different organizational structures 93
9. Competition and Human Needs 93
9.1 Strategies good for profit but that may not be good for human needs 94
9.2 The incentive to supply to over-consumption and human wants 94
9.3 The challenges when human needs are external to the financial systems 95
9.4. Markets and human needs 95
5. PRICING in SUPPLY CHAIN 97
1. Pricing in Supply Chain 97
1.1 Cost-plus pricing 97
1.2 Auctions 98
1.3 Price discrimination 98
1.4 Dynamic pricing 99
2. Auctions 99
2.1 Types of auctions 99
2.2 Private value 100
3. Sealed-bid, Private Value Auction in Class 100
3.1 In-class games 100
4. Modeling and Analysis of Private Value Sealed-Bid Auctions 102
4.1 Modeling of first price auction 102
4.2 Bidder's best responses and utilities when there are two bidders 103
4.3 Best responses and utilities when there are n bidders 106
4.4 Summary basic model with $U(0, 1)$ uniform distribution 109
4.5 Extension to $U(A, B)$ uniform distribution 109

4.6 Vickery second-price sealed-bid private value auction 112
4.7 Summary of sealed-bid private-value auction 114
5. Winner's Curse 114
5.1 Interesting World Cup coverage right bidding in Singapore, 2010 115
6. Price Discrimination 115
6.1 Direct price discrimination with books 116
6.2 Indirect price discrimination 120
7. Dynamic Pricing 120
8. Pricing and Human Needs 120
6. LOGISTICS in SUPPLY CHAIN 121
1. Major Costs in Logistics 121
2. Transportation 122
2.1 Transportation modes 122
2.2 ISO containers 122
2.3 Waterborne freight 122
2.3 Rail 124
2.4 Trucks 125
2.5 Intermodal transportation 128
2.6 Comparison of different modes 128
2.7 Marginal cost of different mode of transportation 129
3. Storage and handling in the supply chain 130
3.1 Function of storage in a supply chain 130
3.2 Considerations in storage strategies 131
3.3 Warehouses, distribution centers, and cross-docks in the supply chain 131
3.4 Marginal cost associated with different handling units 131
4. Logistics Administration 132
5. Logistics Users and Service Providers 133
5.1 Common carriers versus private fleets 134
5.2 Third- party logistics (3PL) 134
5.3 The factors in logistics system design 135
5.4 Network effect in logistics 136

6. The Logistics in the Barilla Case 136
7. Logistics and Human Needs 137
7. ECONOMIES of SCALE in SUPPLY CHAIN 139
1. Economies of Scale in Economics and in the Supply Chain 139
1.1 In economics 139
1.2 Marginal and average costs in the supply chain 140
1.3 Diseconomies of scale 141
2. Scale Effect due to Lower Shared Investment Cost 141
3. Economies of Scale Due to Technology and System Design 145
4. Economies of Scale Due to Batching 150
4.1 Batching to unit load 150
4.2 The dimension of batching 152
5. Economic Order Quantity (EOQ) and Lot-Sizing 152
5.1 EOQ model 152
5.2 Lot-size scale effect from the demand side 157
5.3 Quantity discount 158
6. Economies of Scale in Information Technology 160
7. Economies of Scale in Barilla Case 160
8. Economies of Scale and Supply of Human Needs 160
8.1 The benefits 160
8.2 The drawbacks 160
8. ECONOMIES of SHARING and ECONOMIES of SCOPE 163
1. Optimize the Product Assortments Shared by Variety of Demand 163
1.1 Cost, revenue, and profit with respect to variety 164
1.2 Sharing in a limited product assortment 165
1.3 Choice overload 165
1.4 The inventory-related cost of assortment 166
2. Sharing Resources 168
2.1 Sharing over time and in space 169
2.2 Sharing of equipment or facilities 169
2.3 Sharing in unit load in material handling 170

2.4 Sharing of personnel and cross-training 170
3. Sharing in Input and Work in Process 171
3.1 Select inputs that can be shared by most products 171
3.2 Postponement in differentiation 171
3.3 Modular design 172
4. Economies of Scope 173
4.1 Modeling of economies of scope in supply chain 173
4.2 Computers and humans support flexibility 176
4.3 Characteristics of different supply systems 176
4.4 Omni channel supply chain 178
5. Economies of Scope in Products 179
5.1 Mass customization 179
5.2 Personalization 181
6. Sharing Economy 181
7. Economies of Sharing in Barilla Case 181
8. Economies of Sharing for Human Needs 182
9. UNCERTAINTY, ITS COST AND REDUCTION 183
1. Measures of Uncertainty 183
1.1 Coefficient of variation 183
1.2 Coefficient of variation of some common random variables 185
1.3 Ranges of CVs 186
2. The Cost of Uncertainty 188
2.1 Newsvendor model 189
2.2 Applications of the newsvendor model in the supply chain 193
2.3 Some notation in simple inventory models 193
2.4 Inventory turnover ratio 202
2.5. The strong impact of lead time uncertainty 202
3. Uncertainty Identification and Reduction and Hedging 203
3.1 Reduce apparent uncertainty through analytics 203
3.2 Reduce uncertainty in human processes 205
3.3 Reduce uncertainty in equipment 206

3.4 Reduce uncertainty in input 206
3.5 Buffer demand uncertainty with time, resource, or capacity 206
3.6 Artificial scarcity to speed up the demand 207
4. Uncertainty Reduction Through Pooling 208
4.1 Mathematical basis for the pooling effect 208
4.2 Special case of pooling among identical and uncorrelated random variables 209
4.3 The Impact of Correlations 209
4.4 Dimensions of pooling 210
4.5 Pooling examples 211
4.6 Pooling effect and bullwhip effect 215
5. Variability and Uncertainty in Barilla Case 216
6. Summary of Uncertainty 217
7. The Benefit and Cost of Supply-Chain Uncertainty to Human Needs 217
7.1 The benefits of uncertainty 217
7.2 The cost of uncertainty to human needs 217
10. ECONOMIES of SPEED and NETWORK DESIGN 219
1. The Benefits of Faster Supply 219
1.1 Improve client's supply chain 219
1.2 Increase prices 219
2. Forecast Accuracy Increases with Shorter Forecasting Horizons 219
2.1 Inaccuracies in pooling in time 220
2.2 There is more information in shorter forecasting horizons 220
3. The Components of Order Fulfillment Time 220
3.1 Production time 221
3.2 Production planning time 222
3.3 Shipping time 223
3.4 Shipping planning time 224
4. The Economies of Speed for Human Needs 224
4.1 The benefits of speed 224
4.2 The drawbacks of speed 225
5. Supply-Chain Network Design 225

5.1 The inventory-positioning model 225
5.2 Use inventory positioning model in SC 226
6. Hub-and-Spoke Networks 227
6.1 Point-to-point network 228
6.2 The characteristics of Hub-and-spoke network 228
6.3 Multi-hub system 231
6.4 Revisit to in-house and outsource strategies 232
7. Network Design and Human Needs 232
11. QUANTITATIVE INCENTIVE SUPPLY CHAIN CONTRACTS 233
1. Two-Tier Supply Chain Model 233
1.1 Best location to keep inventory 233
1.2 Risk owner 234
1.3 When the dominant partner does not own the inventory risk 235
2. The Supplier Dominates but the Client Owns Inventory Risk 235
2.1 No incentive contract 235
2.2 Newsvendor model for profit 236
2.3 Optimum profit for each partner without contract 238
2.4 Revenue-sharing contract 240
2.5 Buy-back contract 241
3. Perfect Collaboration or Global Optimum Solution 243
4. The Client Dominates but the Client Owns Inventory Risk 244
4.1 No contract 244
4.2 Pooling effect impact in one-to-many supply tree 245
4.3 Payback contract 246
4.4 Cost-share contract 247
5. Discrete Demand Distribution 248
6. Quantitative Incentives Contract and Human Needs 250
12. SUMMARY 251
1. The Success of Strategies in the Food Supply Chain 251
1.1 Challenges in the food-supply chain 251
1.2 Success of the modern food-supply chain in the United States 252

1.3 Overview of the U.S. food-supply chain 252
1.4 Achievements in farming and the global supply chain 253
1.5 Benefits of food processing with respect to the challenges 253
1.6 Other benefits of food processing 253
1.7 Supply-chain strategies and efficient food-supply chain 254
2. Food Supply and Human Needs 255
2.1 Food desert 255
2.2 Broader impact of low-cost processed food 255
2.3 Organic food-supply chain 256
3. Health, Healthcare, and Human Needs 256
3.1 What can we do? 257
References 258

1. INTRODUCTION TO SUPPLY CHAIN FOR HUMAN NEEDS

This book is based on the materials used in an upper-level college course on supply-chain economics, drawn on experience in teachings operations, logistics and warehousing, and tours and projects over 20 years. The context of the book is the supply chain, which involves manufacturing, logistics, procurements, delivery, and coordination with channel partners. Supply-chain engineering and management have developed many tools to reduce the cost and improve efficiency in the supply chain.

Economics is the study of supply, demand, and markets. It offers useful tools for the analysis of the information asymmetry and for the development of strategies with incentives to reduce moral hazards. This is very useful in the coordination of channel partners. Economics also offers tools to understand competition through pricing, supply quantity, or differentiation. Pricing and differentiation complement the powerful tools developed by supply-chain professionals, with a focus on cost and efficiency. Since the strategies for collaboration and competition depend on the responses from the collaborators or competitors, game theory in economics also provides insights in developing supply-chain strategies. Some of the concepts in economics—such as economies of scale, economies of scope, or economies of sharing—are also useful with supply chains. These terms are widely understood and useful in communications with a broader audience.

A unique feature of this book is the link between the supply chain and human needs—the ultimate objective of human activities in the supply chain! Today, the societies put more and more emphasis on social and environmental impacts. As we will see in many examples in this book, the low-cost, efficient, coordinated, and competitive supply chain serves human needs well—but not always. A supply chain for human needs provides a simple and concrete connection between the supply-chain strategies, economics, social and environmental impacts. The concept of externalities in economics can be used to understand the misalignments between the objectives among finance, society, and the environment.

This book is different from most supply-chain books because it adopts a few economics concepts. This book is different from economics books because it only uses a few tools from economics relevant to certain aspects of the supply chain. This book is different from almost all books, because it links engineering and management to human needs.

This book includes qualitative discussions on the strategies and quantitative models for analysis. The quantitative models require a knowledge of algebra, basic calculus, and statistics. The practitioners who do not have some of this mathematical background can still follow the qualitative strategic discussions.

1. OVERVIEW OF THE SUPPLY CHAIN, ECONOMICS, AND HUMAN NEEDS

The supply chain can be considered a complex network of firms and organizations working together to convert natural resources into the products and services humans need or want, such as water, toilet paper, beverages, food, furniture, cell phones, medical treatment, and classes. The

supply chain in this broad sense is simplified in Figure 1.1. The firms in the supply chain perform various functions, shown in the Supply Chain box. The supply chain takes natural resources from the earth and energy from the sun, and converts them to products and services for human needs. The arrows pointing to the right represent the forward supply chain. The satisfied human beings provide the human resources to perform activities in the supply chain. After use, humans dispose of products back to nature or reintroduce the products to the supply chain for repurposing or recycling. During the process, the supply chain emits wastes back into nature. The disposal and emission are the reverse chain shown by the arrows moving to the left. Therefore, the supply chain is a supply cycle from which human beings draw utility.

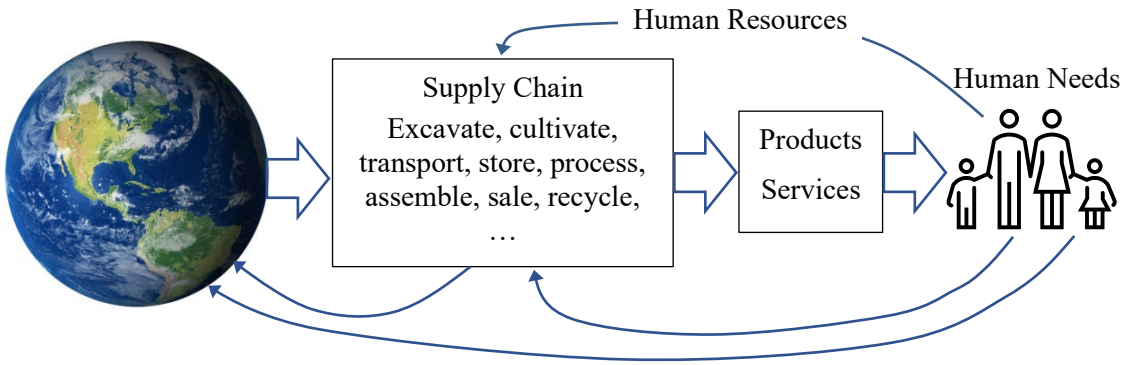


Figure 1.1. Overview of the supply cycle for human needs.

The activities in the supply chain can include excavating raw materials; cultivating crops; raising livestock; processing raw materials; fabricating components; assembling components into modules or products; packaging, handling, and storing materials; and transporting materials. Logistics, which focuses on transportation and storage, is always an important part of the supply chain. All inputs to the supply chain, the raw materials or the energy from the sun, are from natural resources. The results of the supply chain are the products and services we enjoy.

Many activities are involved to convert raw materials—such as oil, sand, and fertile lands—into the products and services humans need. These activities are carried out in many firms along the supply cycle. Internally, the firms must design, plan, analyze, and manage internal activities to reduce cost and improve efficiency. Externally, each firm must coordinate its suppliers and clients. Externally, there are competitors in most markets. They must form strategies to compete effectively. Markets do wonders. It is competition that incentivizes firms to improve efficiency, products, and services.

Human beings purchase the products and services among choices to satisfy their needs and wants. We use the term “needs” to refer to the products and services humans must have to lead a productive life, and the term “wants” to refer to the products and services for additional comfort or enjoyment or a short-term thrill such as a rich cookie after a healthy dinner.

The low-cost, efficient, well-coordinated, and competitive supply chain has achieved abundant supplies for all our needs at low prices in the advanced economies. However, not all is well. Two examples of challenges follow.

In the United States, there are food deserts where the low-income households without cars do not have access to a major grocery store with affordable fresh food. Their nutritional needs are compromised. People lacking balanced nutrition are more likely to have poor health, low productivity, and higher healthcare costs. Poor health compromises the human resources needed in the supply chain. More importantly, children who grow up in such an environment are accustomed to fatty, salty, sweet, strong flavored, and processed food. They will continue to have problems with health, productivity, healthcare costs, and economic wellbeing, leading to many other social problems in the certain communities.

In the United States, the healthcare system does not serve human needs for healthy life well. Healthcare costs are almost twice those in other developed countries. However, the United States has the shortest healthy life expectancies, as published by the World Health Organization (WHO) and shown in Figure 1.2 [1]. More importantly, the United States is the only country with decreasing healthy life expectancies since 2010.

This book will provide insights into the success and challenges of the supply chain for human needs.

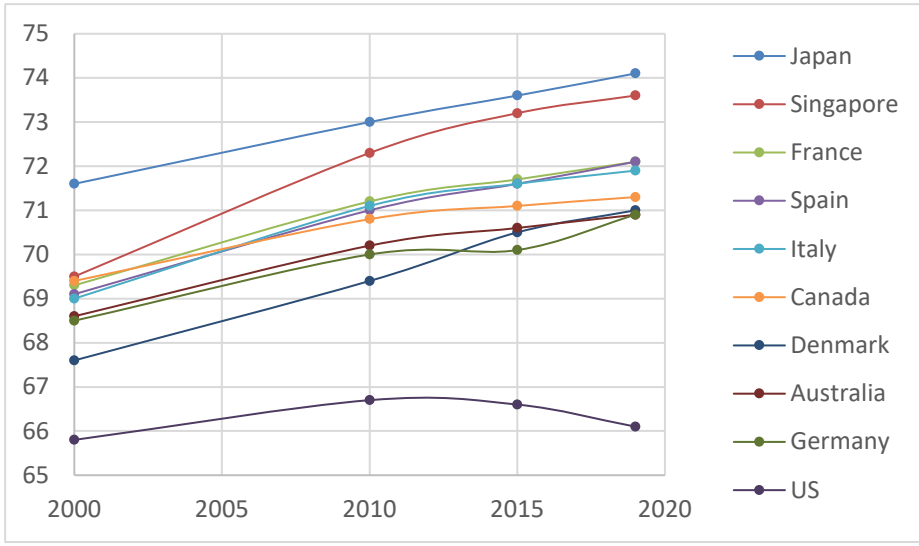


Figure 1.2. Healthy life expectancy at birth (HALE) from 2000 to 2019 by WHO.

2. SUPPLY CHAIN

We will introduce the concepts in the supply chain through the supply of calculators.

2.1 The production chain of a calculator

Many companies supply various calculators, such as Texas Instruments, Hewlett-Packard, or Casio. We call them original equipment manufacturers (OEMs). A calculator consists of many components made of plastics, integrated circuit (IC) chips, metal, and other materials. An OEM will focus on its own core competence, such as design, the production of critical parts, and assembly. It will acquire plastic components, display, connectors, and so forth, from its suppliers. A simplified supply chain shown in Figure 1.3 can be used to illustrate the concepts in this supply chain. Each firm is represented by a box. The acronym in the upper left is the identification of the firm. For this example, OEM refers to the calculator manufacturer and S_{SCR1} is the screen supplier. The value with a dollar sign represents the value to the firm. The two values below represent lead times and lead time variability. More about these definitions and usage later.

The arrows show the flow from the suppliers to clients. The values refer to average lead times and variability. The OEM would source its board BD and screen SCR from its suppliers. These suppliers, in turn, source from their suppliers, and so on, and form a long chain.

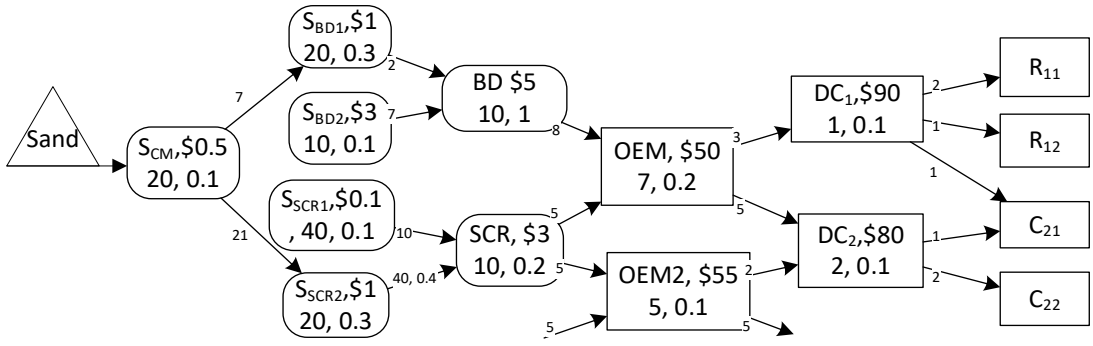


Figure 1.3. Simplified supply chain for a calculator.

Consider the plastic shell of a calculator. The plastic parts supplier would make the components based on the OEM's design using plastic resins in the form of pellets as raw materials. The plastic parts supplier would source the virgin resin or recycled virgin or their mix from chemical suppliers or recyclers, who in term source their raw material from oil companies, recyclers, or other raw material suppliers. The production chain links the raw materials from natural resources to consumer products. The same is true for IC chips, connectors, fasteners, or other subassemblies.

2.2 Distribution of a calculator

After production, the product will go to various distribution channels, shown to right side of OEM in Figure 1.3. The OEM can sell the product directly to end-customer C_{21} via ecommerce and

through their own order distribution center (DC_1) and fleet. This is rare, because OEMs may only want to focus on some aspects of the distribution. The OEM can distribute its products via retailers such as R_{11} or other distributors or wholesalers not shown. They can also send their product to others distribution centers such as DC_2. Some OEMs can also sell their products directly to customers through their own distribution center (DC_2) using the Amazon Marketplace platform. Nowadays, there are many possible combinations of distribution channels with different arrangements of storage, transportation, marketing, ownership, and finances to take advantage of the scale and competencies of firms in each sector in the distribution system. E-commerce has already changed distribution a lot, and will continue to do so for the years to come.

At the end of use, the calculators can be reused by others, recycled for raw materials, or disposed of in landfill, completing the supply cycle.

2.3 Supply chain defined by CSCMP

The Council of Supply Chain Management Professionals (CSCMP) defines the supply chain in this way:

"Supply chain management encompasses the planning and management of all activities involved in sourcing and procurement, conversion, and all logistics management activities. Importantly, it also includes coordination and collaboration with channel partners, which can be suppliers, intermediaries, third party service providers, and customers. In essence, supply chain management integrates supply and demand management within and across companies."

The first part of the definition is the planning and management of sourcing from suppliers; the conversion process, such as fabrication or assembly; and logistics functions such as transportation and inventory in the inbound direction from the suppliers and the outbound direction to the clients.

The second part of the definition is the coordination and collaboration with channel partners. In addition to the suppliers and clients, a firm may also need to coordinate with the intermediaries, such as consulting firms and third-party service providers, such as third-party logistic providers (3PLs). The coordination and collaboration can involve communication, contracting, mutual assistance, and so on.

Together, the planning, coordination, and collaboration connect the supply from suppliers, the firm's value-added services, and the demand from customers.

Supply-chain management and engineering offer many tools for planning, coordination, and collaboration, such as (1) forecasting the demand; (2) designing, planning, and execution of transportation, inventory, warehouses, and procurement; (3) customer relationship management (CRM); and (4) the analysis of in-house supply or outsourcing of certain functions. The objectives of the tools in supply-chain management are normally cost reduction or efficiency improvement, as in fewer resources and faster and higher-quality service, which often also convert to financial measurement.

3. MICROECONOMICS IN THE SUPPLY CHAIN

Microeconomics defined in Intermediate Microeconomics (Nicholson and Christopher Snyder, 2015) is:

"Microeconomics is the study of the economic choices individuals and firms make and how these choices create markets."

This definition is directly related to the supply chain. The individuals' needs, choices, or self-interests generate the demands. In the calculator example, the students draw utilities from the calculators to learn and to take SAT. The OEMs respond by supplying the product consumer need. The OEM will source the supplies from its suppliers and collaborate with the distributors to distribute the products to the consumer. The OEMs supplies products to satisfy the demand, which in turn generates demand for its own suppliers.

3.1 Collaboration

All supply-chain professionals understand that they must collaborate with their channel partners such as suppliers, clients, and service providers. The supply chain offers certain tools for the coordination and collaboration. However, direct contention exists between the supplier and the client. For example, the profit of the supplier increases with the prices, while the profit of the client decreases with the prices. Economics offers additional tools to develop strategies to align the objectives of the supplier and the client. For example, the retailer owning the distribution center DC_2 in Figure 1.3, is a client for the OEM and OEM2, a competitor that also supplies calculators. The objective for the calculator OEM is to maximize profit selling its calculators. However, the DC_2's objective is to maximize the profits among all its suppliers. When there is contention, DC_2 may not always prioritize for OEM. The inverse is also true for a supplier to multiple clients. The challenge in assigning incentives is the information asymmetry.

Each firm keeps a lot of information regarding its processes, plans, personnel, customers, suppliers, clients, finance, contracts with their partners, and so on. Some are proprietary and some can be shared with the public. However, it is often difficult to determine in an optimum way. Most firms tend to take a conservative approach with a tendency to keep the information private. However, there often exists information that if shared with channel partners, can benefit one or both partners. For example, the daily sales at stores can be very valuable to the suppliers to plan and manage their supply chain. Since the supplier and the store are not direct competitors, proper sharing does not constitute a threat to the store. The supplier may be willing to share some benefits gained from the information to the store. This can create a win-win situation for both the supplier and the retailer.

The principal-agent models in economics offer models for information asymmetry, adverse selection, and moral hazard. Adverse selection is the selective revealing of information, as when the retailer does not reveal its relationship with other OEMs or does not reveal past failures. Moral hazard refers to the choice to hurt others without hurting the choice maker, such as the retailer pushing the sales of other OEMs' calculators, or wasting food at an "all you can eat" buffet. The principal-agent models also provide tools for the use of incentives to reduce adverse selection and moral hazard.

3.2 Competition

Figure 1.3 showed a competitor OEM2 that also supplies calculators. OEM and OEM2 must compete in the market to make profits by setting supply quantity or setting the price. The market will balance the supply, the demand, and the price to reach the equilibrium. In each market, the

pricing signal helps the suppliers to develop their strategies. Nobel laureate Fredrick Hayek also observed that is it the prices along the chain that link the supply and demands from raw materials in natural resources all the way to human needs and wants. The self-interests of the firms and the consumers, described as invisible hands by Adam Smith, linked these different self-interests and prices to promote the efficient market.

The cost reduction and efficiency improvements in supply-chain engineering and management are major contributors to be competitive. The lower cost increases profit as the same price determined by the market. Better product and better services can allow higher prices. Microeconomics offers additional models for competition through pricing, supply quantity, or product differentiation through design, quality, or customer service.

In competition, the results of a strategy depend on the strategies of competitors. Game theory offers insights into developing best strategies for the firms.

4. SUPPLY TO HUMAN NEEDS

We use "needs" to refer the things and services a human being must have to lead a productive life in modern society. Humans have bodies. Their bodies need food, clean water, fresh air, basic sanitation, shelter, health, livelihood, and safety. In modern society, the livelihood needs education and healthcare. Although rapid technological advancement has made tremendous changes in the supply of these, our bodily needs do not change quickly.

The stay-at-home orders in the Covid-19 pandemic shed light on what's needed. The "essential workers" must work to supply foods, utilities, delivery services, medical services, and so on. We can use "wants" to refer the things for additional comfort or desire not necessary to lead a healthy and sustainable life, such as excess food, a 72-degree constant temperature, substance abuse (opioids), ripped jeans, tattoos, or comfort instead of needed exercises. When we are thirsty, we need and want water. When we are stressed or bored, we may want rich snacks, but we do not need them. We need healthy food, but sometimes we want tasty food or excess food. However, there is no well-accepted definition of human needs or wants.

Most people know Maslow's hierarchy of human needs, shown in Figure 1.4. Although the upper-level needs can be controversial, no one would question the importance of the needs in the

Physiological and Safety levels. These are necessary for a productive life in the modern society.

Most of staple products and services from the supply chains are related to the physiological and safety needs. We can also look at our expenditures. The results from U.S. Bureau of Labor Statistics, "Consumer Expenditure Surveys, 2018" [2] shows how consumers allocate their financial resources related to their needs and wants from the supply chain. Table 1.1 lists the average household expenditure percentages in order of essentiality to our lives.



Figure 1.4. Maslow's hierarchy of human needs.

Table 1.1. Consumer expenditure with respect to human needs.

Human Needs listed in the Order of Essentiality	**Expenditure (2018, %)**	**Comments**
Air	0.0	
Water (public services)	1.0	
Food	12.9	7.3 home, 5.6 away from home
Sanitation		Mixed in different categories
Housing	32.8	Shelter 19.2, utilities 5.6, furnishing 3.3, etc.
Livelihood: Transportation	15.9	Vehicle purchase 6.5, fuel 3.4, etc.
Livelihood: Healthcare	8.1	Insurance 5.6, (17.7 GDP from cms.gov, including other sources such as tax).
Security: Insurance	23.9	Personal and pensions 11.9, life and other 0.8, pension and Social Security 11.2
Livelihood: Education	2.3	
Off springs		

From this list, we notice that we do not spend much on the most important human needs, such as fresh air and fresh water or utilities. We spend about 12.9% on food, 32.8% on shelter, 15.9% on transportation, 8.1% on healthcare, 23.9 on insurance or safety, and 2.3% on education. Please note that this does not include large amounts of money consumers pay through federal, state, local, sales, and other taxes. Federal revenue and expenditures in 2019, the year before pandemic, are shown in Table 1.2. The expenditures are almost all for basic human needs. Medicare and Medicaid expenditure is almost one quarter of the entire expenditure!

Table 1.2. Federal tax revenue and relevant expenditures in 2019.

Federal Revenue, $3,500 billion		
Category	Amt. (billion)	%
Individual income tax	1,700	48.6
Payroll tax (Social Security, Medicare)	1,200	34.3
Corporate income tax	230	6.6
Estate, customs, etc.	271	7.7

Federal Expenditure, $4,400 billion		
Category	Amt. (billion)	%
Social Security	1,000	22.7
Defense	676	15.4
Medicare	644	14.6
Medicaid	409	9.3
Transportation, education, veterans, housing, etc.	661	15.0

5. THE OUTCOME OF THE SUPPLY CHAIN TO HUMAN NEEDS

5.1 Supply chain achievements

The cost and efficiency-focused supply chain has achieved great success. People in the developed countries enjoy an abundance of products and services for their needs when they needed them. For most people, we only spend a small fraction of our income on what we need. The total products and services for basic human needs are more than sufficient for the entire population in the developed countries for many years. A supply-chain professional today can live a life more comfortable than even the members of royal families when Adam Smith wrote *The Wealth of Nations*. Many people can also enjoy lifestyle products and services such as special drinks, food, diet, fitness, style, decorations, luxury transportation, and luxury houses that are beyond basic needs. Even the widespread business closings during the Covid-19 pandemic did not disrupt the supply chain to the level of significantly impacting people's lives. We should celebrate these achievements.

The great achievements also transform some of the basic economic dynamics.

5.2 Abundance and scarcity

Before the industrial revolution, the products produced supplied by the supply chain were insufficient for the physiological and safety needs for the population. Life was harsh, with much shorter healthy life expectancies. Food, shelter, sanitation, and healthcare services were scarce. On the other hand, the natural resources, such as fresh air, fresh water, minerals, plants, and land were in abundance.

Nowadays, the supply capacity in the developed economies exceeds what people need in many aspects, such as food, shelter, sanitation, and basic healthcare services. The supply capacity is sufficient or more for all the population in the developed countries. However, some of the natural resources, such as fresh air, fresh water, certain minerals, and land become scarce.

One of the major assumptions in economics is supply scarcity. However, the fact today in many areas is supply abundance. The suppliers invest to supply more of what we want, since the needs market is saturated.

There are other scarcities: time to sleep, to be with loved ones, to savor, or to think through the overflow of information. The time scarcity is related to information abundance and entertainment abundance on smart devices. Time scarcity has contributed to the rapid increase in anxiety and depression and lead to a scarcity of mental health services.

5.3 The assumption of rational choice

Another major assumption in economics is that people make rational choices for their own interest. As people become more and more affluent, the penalty for making wrong choices become lower and lower, since people have more disposable income. In addition, people have the luxury to make choices based on their short-term emotions or desires instead of their needs or long-term benefits. The suppliers, equipped with artificial intelligence and unprecedented access to personal information, know how to push people's emotional button to direct people's choices, often away from their rational choices.

5.4 Human needs and human resources

The satisfied humans provide human resources to perform the activities in the supply chain. Therefore, it is to society's benefits to satisfy human needs to ensure sufficient human resources to support the activities in the supply chain.

Jobs becomes more and more specialized and more efficient. For example, smart phones are very complicated devices. However, their supply chain is so efficient that it does not need many people to supply the entire world. Many jobs today require more and more training. There are shortages in many professions that require certain knowledge. Education becomes a basic human need.

People deprived of basic needs such as nutrition, care, safety, education, early in life before they make their own decisions will have low chance to become human resources or a productive member in the society. In fact, they are likely to become problems to the society that deprive other people's needs, including their own off springs.

5.5 Pricing signals and human needs

Hayek's pricing signals reflect the exchange values in the markets. Because of supply abundance and low cost of irrational choices, the exchange value in the financial terms may not always align with the value of human needs.

Clean air is arguably the most valuable human need. Yet, it is basically free. The United States is blessed with low population density and cleaner industry because many air-polluting industries are offshored. Most people do not suffer from poor air quality. However, there are still communities and professions that are exposed to unhealthy air, such as neighborhoods near polluting industries or people working in coal mines, shale fracking fields, or with gas-powered leaf blowers and other garden equipment.

Air pollution is much more severe in certain parts of the world such as Beijing, New Delhi, or Lima. You can search NASA's data on carbon dioxide motion around the globe. You will notice that no matter which parts of the globe we live in, the air is circulated by the atmosphere. The air pollution in one region can move quickly to other parts. The problem cannot be solved by a single country. Clean air is a public good for the world that cannot be easily controlled or priced.

Fresh water is similar to fresh air: low in exchange value but a very important human need. Even in the United States, the water problem in Flint, Michigan, exposed in 2015, is still not resolved by 2021! The United States is blessed with fresh water in most parts of this large country. However, freshwater flow through many countries in Africa and South Asia poses many health and other problems for the population.

If the market sets low prices on things important to human needs, and high prices on human wants, then the financial objectives may not align with the human needs objectives. We must identify these challenges and seek solutions.

5.6 Economic externalities

Externalities are an important economics concept that captures the impacts of economic activities external to the economic systems. The objectives in economics are utilities. Since people can draw utilities broadly in human, social, or environmental measures, human needs can be included in the economic utilities. It is more accurate to consider externalities as the impact of supply-chain activities external to the financial systems. Hayek's pricing signals only reflect the financial values. Many important human needs, and social and environmental utilities, are outside of financial systems in a country, such as clean air or consumer health.

6. EXAMPLES OF MISALIGNMENT BETWEEN HUMAN NEEDS AND PRICING

Because of misalignment between the financial exchange values and human needs values, the very successful supply chain can cause problems. Here are a few examples.

6.1 Food deserts

We already introduced food deserts earlier in this chapter. We can now use cost, efficiency, competition, abundance, rationality, and personal choices to give some qualitative explanations. In the food desert, there are packaged foods available in food marts, convenient stores, gas stations, and prepared food in fast-food restaurants. However, there is no fresh food typically sold at major grocery stores. Below are some economic factors for food deserts.

1. Foods are from farm products. Farm products are seasonal, with a short shelf-life. In addition, fresh fruit and vegetables are hard to handle, with a low yield. The density in transport, storage, and display in the store is low. They also need temperature and environmental control. Even with all the strategies developed by the supply-chain professionals, the cost of supplying fresh food is still higher than low-cost processed food. This is because the engineers have developed many effective processes to remove the challenges associated with supplying fresh foods.
2. The margin in grocery retail is low. The supermarkets carrying the fresh produce must also carry many other products to attract many customers, in order to be profitable. There must be a large customer base to justify a major supermarket. On the other hand, the food mart or convenience store can be small with low revenues, carrying only low-cost processed foods.
3. Food processing also provides an opportunity to enhance the taste. Due to the abundance of fat, sweetener, and other additives, the supplier can create tasty foods targeted to certain populations at very low cost.
4. Processed foods are also more convenient in consumption: open, mix, heat, and eat. Fresh food, on the other hand, requires washing, cutting, and cooking. These take a lot of time, requiring facilities and skills.

Together, these characteristics give processed food advantages in price, accessibility, taste, and convenience over fresh foods. However, the low-price processed foods do not satisfy human health needs. They contain excess salt, sugar, and other additives humans do not need. Excessive consumption of processed food will lead to poor health, loss of productivity, higher healthcare cost, and social problems that can hurt the society. However, these problems are external to the financial systems of the firms in the food supply chain. The same is true with the fast food and family restaurants.

In fact, the over-consumption of processed foods from the cost-driven supply chain is also a contributing factor for the short and decreasing healthy life expectancy shown in Figure 1.2.

6.2 High cost and low performance healthcare supply chain in the United States

The healthcare cost in the United States in 2018 was 17.7% of the GDP, almost twice that of the other developed economies. However, in one of the most important measures of health, healthy life expectancy reported by the World Health Organization (WHO) and shown in Figure 1.2, the United States is low compared to other developed countries, and is decreasing! The definition of "healthy" is a complex set of parameters. It roughly refers to the condition of the person still being reasonably self-sufficient. Interestingly, United States has the largest gap from healthy life expectancy to life expectancy, and with the fastest increase! This means that the healthcare supply chain is good at keeping people alive but not healthy! One of my students said that the United States has a sick-care system, not a healthcare system. No wonder 65% of the registered voters in the 2020 election considered healthcare "very important" (Pew Research Center [3]).

6.2.1 Who pay how much for the medical bills?

Figure 1.5 shows the sources of expenditures from 1980 to 2018 published by the Centers for Medicare & Medicaid Services [4]. Note that the first four points are not per year. The four solid lines are explained below:

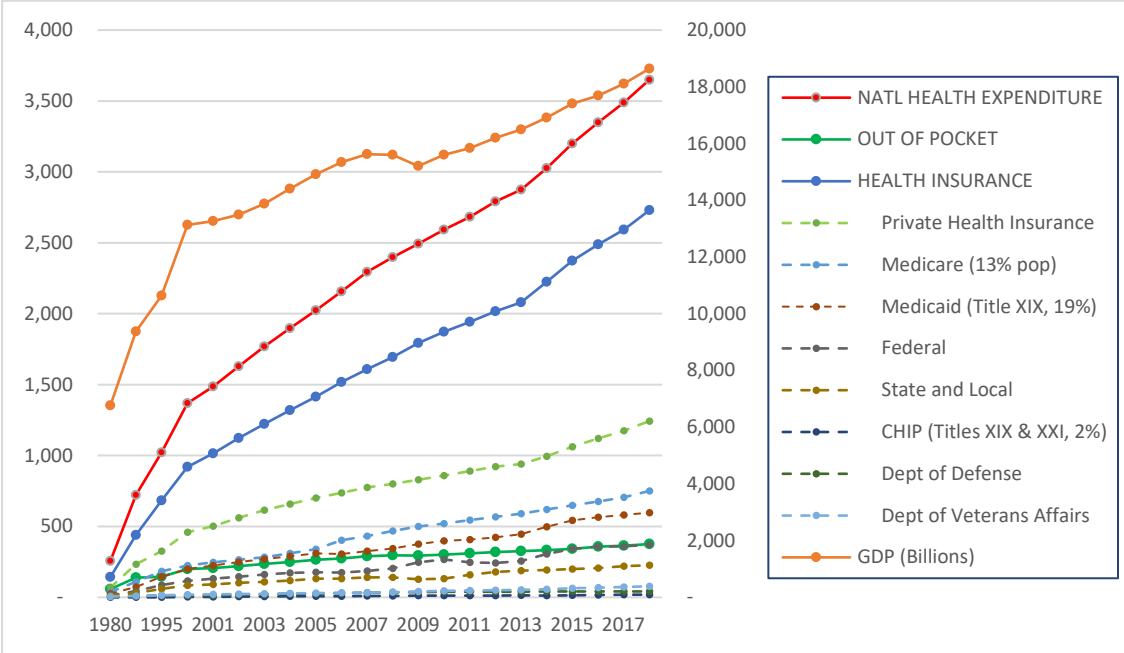


Figure 1.5. Source of expenditures of medical services based on the survey by Centers for Medicare & Medicaid Services.

1. The top solid line is GDP. Its annual rate of increase averaged 1.96% in 18 years.
2. The second-highest solid line is the total NATIONAL HEALTH EXPENDITURE. Its annual rate of increase was 5.6%, 2.86 times the growth of the GDP.
3. The third-highest solid line is the total expenditures from all HEALTH INSURANCEs, including both private health insurance and government insurances such as Medicare and Medicaid. The annual increase rate is 6.23%, 3.17 times the GDP growth rate. The sum of expenditure from Medicare and Medicate is higher than all private insurance combined.
4. The fourth solid line is the OUT OF POCKET. The annual rate of increase is 3.60%, 1.84 times that of the GDP.

All the other dashed lines are government expenditures. Since the sources of government expenditures are from taxes. Table 1.2 shows 85% of federal taxes are from income tax and payroll tax. The private insurance firms are for-profit organizations. The premiums are paid for by the consumers or subsidized by the employers; either way, they are ultimately from the consumers. Therefore, almost all of the healthcare expenditures, in the amount of 17.7% of the GDP, are ultimately paid for by consumers, some in the form of out-of-pocket co-pay, some

through private health insurance, and some by government programs from the mandatory taxes. Since the consumers pay for the insurance premiums, government taxes, and out of packet co-pay, all 17.7% of the GDP is ultimately paid for by the consumers.

6.2.2 Characteristics of the healthcare supply chain

Healthcare is very different from the consumer goods supply chain in many ways. Here are some characteristics of the healthcare supply chain.

1. Healthcare needs are highly uncertain. Insurance is necessary to hedge against the catastrophic risk individuals face. However, insurance consumes resources without direct supply of healthcare.
2. Demand is insensitive or inelastic to price. Patient will pay even when it is too expensive.
3. Government is always part of the system, for many reasons. Figure 1.5 shows that 57% of the expenditures are paid for by some government programs. Government involvements weaken the market forces and increase opportunities for moral hazards.
4. The size and monopolistic nature of the government program allow the government to negotiate payments at below the cost, leaving the private firm no choice but to balance the finances in other ways and complicating the problem.
5. Information asymmetry
    a. Medical knowledge and pricing are complex. The doctors, hospitals, and insurance companies have much more information than the patients.
    b. Patient information is private, and makes information transparency difficult.
    c. There is opportunity for vendor-induced demand (doctors are often encouraged by the management or consultant).
    d. Patient may not reveal preexisting conditions and waste resources without suffering direct penalty.
    e. Contracts between doctors, hospitals, and insurance companies are hidden from the patients.
6. It is difficult to classify the large set of possible medical services. The coding is so complex, there are master's degrees offered to train professionals to code the treatments and procedures. The management often encourages coding optimization with financial measures.
7. Infectious disease can impact others. It may be more efficient to subsidize vaccination, face masks, identifying and treating patients for the public good. Covid-19 is one example. Other examples include measles and polio.
8. Human needs include both health and care. The medical system is designed for care. Therefore, there is objective misalignment between the healthcare supply chain and human needs, see Figure 3.5 in Chapter 3.

Combined, the healthcare system presents more challenges between the financial measures and human needs measures.

6.2.3 Complexity of the healthcare supply chain in the United States

Individually, the hospitals, doctors' offices, and insurance companies all work hard to reduce cost, improve efficiency, increase resource utilization, and speed up the processes. They develop

automated or other assistance systems. These are very helpful for cost reduction and service improvement.

However, there are many players in the system: doctor's offices, hospitals, pharmacies, and insurance companies. Behind the scene, there are many other organizations in the healthcare supply chain. The major ones are shown in Figure 1.6.

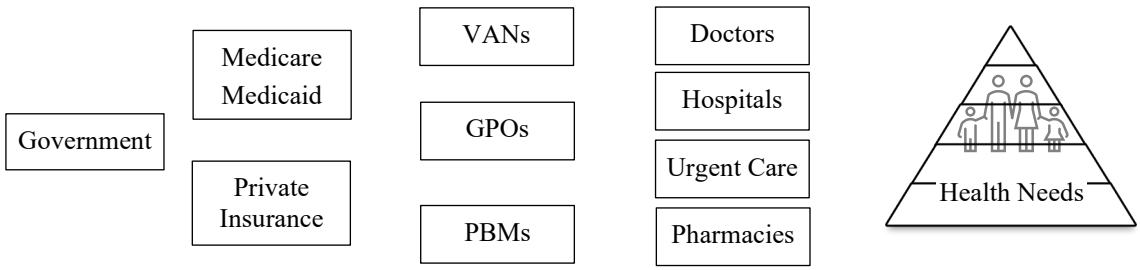


Figure 1.6. The major organizations in the healthcare supply chain.

The Group Purchasing Organizations (GPOs) organize hospitals to increase negotiation power with manufacturers, distributors, and other vendors. The Value-Added Networks (VANs) provide secure data exchanges between organizations via Electronic Data Interchange (EDI). Since healthcare data is complicated, involving regulations, privacy, and complex medical codes, VANs provide additional services in the healthcare supply chain. The Pharmacy Benefit Managers (PBMs) act as middlemen between insurance companies and pharmaceutical companies and pharmacies. These organizations have complex contractual relationships hidden from the patients.

For someone with insurance, there are at least two payers: the patient and the insurance. Sometimes, there are multiple insurance companies involved. There is no incentive to reduce the total cost, ultimately paid for by the patient or other people in the country. The contracts and negotiations between the parties consume many resources without providing any direct service to the patients. In recent years, an interesting phenomenon has occurred. The hospital or doctor's office may tell the patient, if they pay out of pocket, the charge is X. If they file for insurance, the patient will be charged more than X. This reveals some of the problems in the system!

A project team in ISyE 4301 Supply Chain Economics in the fall of 2020 studied healthcare systems in Canada, England, Germany, Singapore and the United States [5]. They compared expenditure as percentage of GDP against patient wait times and doctor-to-patient ratios, shown in Figure 1.7. The expenditures in Canada, England, and Germany are similar, and all much lower than the United States. Among these three countries, Germany has the shortest wait times and more doctors per patient. England has shorter wait times than Canada. Singapore has the lowest expenditures, fewest doctors, and shortest wait times! More importantly, Figure 1.2 shows that Singapore also has the second longest healthy life expectancy!

It is worthwhile to see how the Singaporean healthcare-supply chain works.

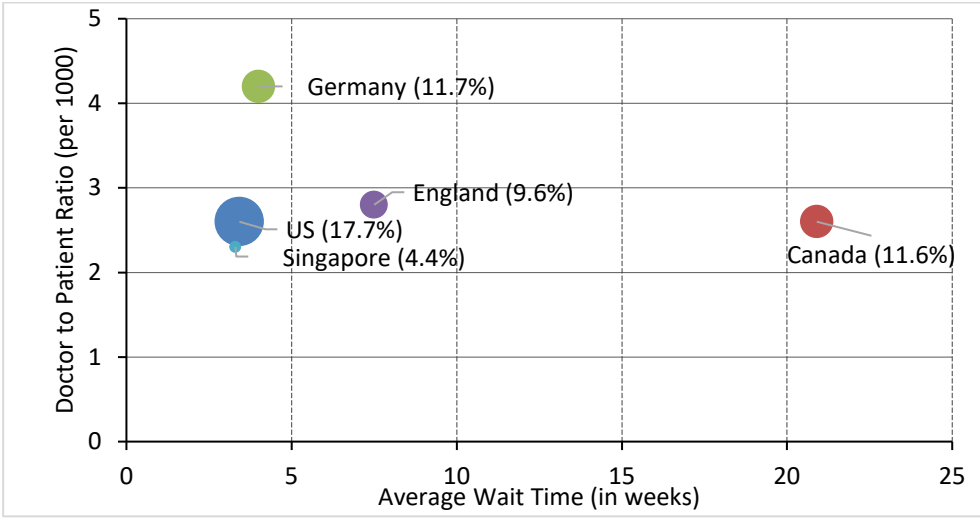


Figure 1.7. Healthcare expenditure as a percentage of GDP against patient wait times and doctor-to-patient ratio.

6.2.4 Healthcare in Singapore

The Singaporean system has the following characteristics:

1. Pricing is more transparent (https://www.moh.gov.sg/resources-statistics). For example, you can find the median prices for patients with or without government subsidy of removal of the appendix in different hospital bed selections. Pricing transparency is a necessary condition for the patent to choose.
2. Wait times are published. This is another useful metric helping patients to choose.
3. There is a cost for any service except for people with extreme needs. This reduces the moral hazard of over-consumption by the patient.
4. There is a mandatory Medisave program through payroll deductions. The patient can use this money to pay for medical services.
5. Government subsidies for services depend on the level of necessity: the higher the necessity of the service, the more the subsidy.
6. There is government-run insurance, MediShield, which covers the more expensive and necessary services.
7. There are private health insurances for people who want additional services beyond basic coverage.
8. The structure and payment schemes are predetermined. No negotiations are needed after treatment. This reduces the administrative costs.
9. There are fewer players in the system, such as GPOs and PBMs.
10. Wide participation in the use of national electronic health records.

All of these help to align the financial measures with human health measures.

6.2.5 Some reasons for high cost and low health performance in the United States

First of all, there is no transparency in cost and service. The hospitals use "charge masters" to determine how much to charge based on insurance plans, government programs, patient ability to pay, and so on. For the drugs, there are "list" prices, but the actual prices can differ significantly based on the discounts and rebates among different players, and other information hidden from the public.

When patients adversely select to reveal only good information, the insurance company must protect themselves and raise prices for everyone. The providers can also adversely to share certain information with the patient and insurance companies to gain advantage.

Senior design teams working with hospitals find that hospitals through away packages of materials when patients check out. They found that the insurance has paid for the material! This is just one small tip of the iceberg in the moral hazard due to many complex contracts between collaborating pairs of organizations. Patients can over-consume services or through away medications when the expensive drugs are paid for by the insurance.

The ultimate loss of health, productivity, and excessive pay by the patients is external to the service provider's accounting systems. There is no direct incentive to reduce the cost for the real human needs: health.

6.2.6 Summary of the challenges in healthcare supply chain

Through lenses of who is paying how much, characteristics of the healthcare supply chain, the complexity of the healthcare supply chain in the United States, the comparison with other countries, the information asymmetry and invisibility, and adverse selection and moral hazard among the complex network of many players, we can understand more about the challenges in the healthcare supply chain in United States.

6.3 Other examples of misaligned supply chains and human needs

6.3.1 N95 facemasks

One of the challenges during the pandemic is N95 facemasks. These masks are critical protective gear for health workers' livelihood. However, the United States struggles to supply sufficient N95 facemasks even after a year and a half. The N95 is a low-value product. The cost-driven supply chain moved the supply offshore to reduce the cost. However, this low-cost, low-value product is vital for human health!

6.3.2 Essential workers without essential compensation

In the United States, worker's pay is determined by the firm and the market. During the pandemic, many "essential" workers such as online fulfillment center order pickers, meat packers, delivery and ambulance drivers, and crop pickers are low-wage earners. Many do not have health insurance or have insufficient health insurance. The high out-of-pocket medical payments in the United States can cause significant hardship to this vulnerable population. If their work is essential, why can't their compensation support the essential livelihood? Many of them live in higher-density accommodation with other essential workers and so are more likely to be exposed to the virus.

How can we, as a society, ensure that the human needs of essential workers are satisfied?

6.3.3 Challenges for the lithium battery supply chain

Many experts predict a potential shortage in the lithium supply due to the rapid demand increase in batteries. The shortage is not the lithium itself, but the processing capacity. The cost of processing is low by the current suppliers. However, the costs to develop new capacity in other known sources are high. The quarterly financial reporting system would not incentivize long-term investment.

6.3.4 Computer chip shortage

The supply chain for computer chips is long and complicated across different countries. The pandemic reduced the capacity of the chip supply chain but increased the demand for chips for smart devices. In 2021, the world is experiencing significant chip shortage. It has impacted even car supply, since modern cars need many chips to run. The chip and N95 shortage showed that the financial driven supply chain can lack the resilience, a hot topic nowadays.

7. THE COST OF UNSATISFIED HUMAN NEEDS

This chapter starts with Figure 1.1 to show that the goal of the supply chain is to supply for human needs. The figure on the book cover, repeated in Figure 1.8, place the human needs are inside a pyramid described by Maslow. People with satisfied needs enjoy life and provide human resources to the supply chain.

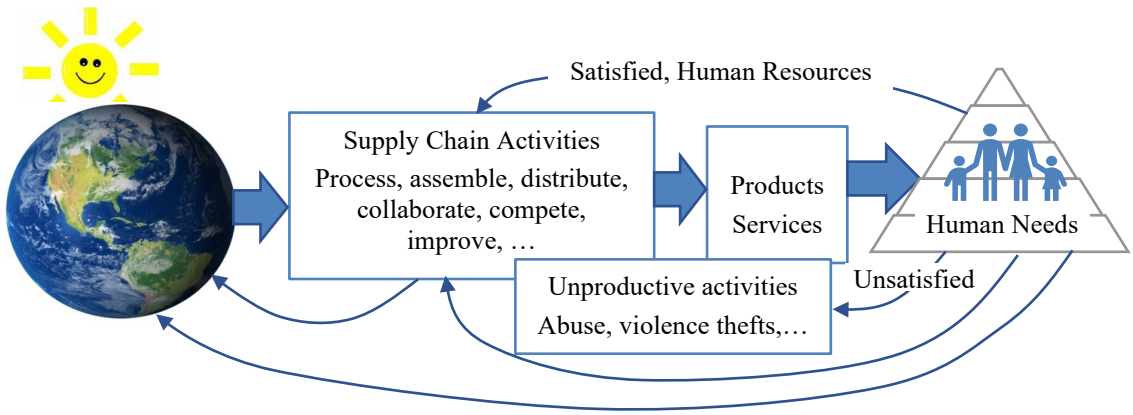


Figure 1.8 The cost and benefits of satisfied and unsatisfied human needs.

However, not everyone's needs are satisfied. For example, an infant has many needs that must be provided by the others: shelter, nutrition, love, safety, guidance, and opportunity to explore and to learn. Most people in US take these for granted. However, many grow up in an environment with food insecurity, poor nutrition, abusive caregivers and dangerous neighborhood. During this time, an infant has to gradually learn to satisfy his/her own needs and learn the life skills and to make decisions. It takes many years before he/she can earn his/her own livelihood. However, youngsters malnourished, abused, scared and deprived of good guidance and good education will

face much more challenges on the way to earn his/her own livelihoods. Some may never develop the skills and abilities to become a productive citizen. To make a living, some may engage in unproductive activities such as substance abuse, theft, or illegal trade. These activities add cost to themselves and to others in the form of insurance, damages and personal safety. Such behavior can influence the next generation, as discussed in the food desert section. These activities also increase the cost for jails, legal service and law enforcement. These costs of the unsatisfied human needs should also be considered by the supply chain professionals and everyone in the society.

8. THE CHANGE

More and more people realize the importance of human, social, and environmental issues. The human, social and environmental issues are closely related to human needs.

Many people and firms have recognized that financial measures alone may not address some of the most important social and environmental issues. In the food supply chain, the movements of local food, slow food, organic food, pastured livestock, urban gardening, and community-supported agriculture (CSA) have emerged. However, without the scale, pesticides, herbicides, feedlots, GMO seeds, preservatives, and processing, the healthier and more sustainable food costs more in production, transportation, storage and retail. To increase the consumption of pastured, nutritious, and fresh food for the population, there is a need for change on the demand side, with people putting higher value on human health. If more people choose the healthy foods, the firms will invest and the market will deliver. Food better for human needs will then enjoy the economies of scale, as well as many other modern technologies, such as connected and smarter communities.

Educators, corporations, and government put more and more emphasize on the social and environmental issues. Most organizations today emphasize the importance of triple bottom line: social, environmental and economic, Or People, profit and planet.

8.1 Human needs and triple bottom line in education

In engineering education, an accreditation agency ABET listed seven student outcomes (SO). Among them, two are about these broader perspectives:

SO 2: An ability to apply engineering design to produce solutions that meet specified needs with consideration for public health, safety, and welfare, as well as global, cultural, social, environmental, and economic factors.

SO 4: An ability to recognize ethical and professional responsibilities in engineering situations and make informed judgments, which must consider the impact of engineering solutions in global, economic, environmental, and societal contexts.

All engineering programs must demonstrate that students are trained to consider these complex issues. However, these issues are overwhelmingly complex for college age students, even for most faculty.

Human needs offers the first principles to link to the complex social, environmental and economic through concrete considerations. We all need air, water, food, safety, livelihood, safety, respect and hope. In modern society, livelihood needs technical and interpersonal skills from extensive education from caregivers and from society. Modern livelihoods also need transportation,

occupational safety, etc. However, the markets often incentivize the supply to human wants because yesterday's luxury become today's necessity for the affluence.

8.2 Corporations

Corporations have recognized the importance of satisfying the human needs of the essential workers to ensure sufficient human resources in their businesses. The mission of most supply-chain firms nowadays almost always includes social and environmental responsibilities, which link the product and services the firm provides with human needs. Not paying sufficient attention can lead to disasters in public relations, fines, and tarnishing a firm's reputation and sometimes the bottom line.

Many supply-chain professionals call for a "stress" test for the essential supplies during special situations such as a pandemic. This will incentivize the corporations to consider long-term impact of supply-chain strategies.

8.3 Government

Many government programs, NGOs, and firms tackle various problems when the financial measures do not align with human needs measures, such as food deserts, high cost in healthcare, the environment, and others.

For example, in healthcare, a new federal rule took effect on January 1, 2021. It requires hospitals to reveal their charges. The *Wall Street Journal* (WSJ) found that the prices for several common medical procedures can vary more than five times [6]. The charges are related to the negotiated deals between the hospitals and insurance companies and are hidden from the consumer.

8.4 Government or market

Markets work wonders if human needs drive the market. However, some human needs, such as air, fresh water, sustainable environment, long term impact, and health are often external to markets. For such the solution may involve the government because the government must serve the public or common good. However, government may not be efficient, especially if the problems and solutions are complicated. Another possibility is for the government to focus on the design of the markets and let market to perform the tasks to satisfy the human needs. For example, healthcare works very differently among the countries mentioned above.

9. ORGANIZATION OF THE BOOK

9.1 Structure of the chapters

This book discusses strategies and models in supply chains and microeconomics. We will start with the big-picture economic concepts on collaborations, competition, and pricing. We will then discuss the specific cost and efficiency concepts in the supply chain. In each chapter, we will also discuss the alignment of the finance-driven strategy and human needs. Specifically, the contents are divided into the chapters with the numbers below.

2. Collaboration in supply chain. We will discuss the importance and challenges in collaboration such as objective misalignment and information asymmetry through a case study. We will then refer to an article to discuss various strategies and incentives to improve the collaborations.

3. Principal-agent model. Principal-agent models are developed to study the information asymmetry between partners, and the design and analysis of strategies to align the incentives to achieve synergy. The principal-agent model allows the supply-chain partners to design and analyze strategies to reduce adverse selection and moral hazard to achieve common goals.
4. Competition. We will first discuss the two classic models: (1) setting supply quantity to compete following the Cournot model and (2) setting price to compete in the Bertrand model. We will link competition with the game theory. We will then discuss their extensions on more suppliers and different costs. We will then discuss the differentiations and the Bertrand model with price differentiation. Finally, we will discuss how to apply these theoretical results to develop practical competition strategies.
5. Pricing in supply chains. We will first discuss some common pricing strategies, including the pricing in the competition. We will then discuss a specific form of an auction: a sealed-bid private value auction, which is common between the supply-chain partners. We will finally discuss price discrimination, meaning setting different prices for different clients, such as the "charge master" in hospitals.
6. Logistics in supply chains. There are two major components in logistics: transportation and inventory. The specialization and globalization make transportation necessary to relay the physical materials from the many suppliers to many clients along the supply chain. In order to have what is needed at the right location at the right time with the right quantity, inventory is necessary. In this chapter, we will discuss various issues associated with transportation and inventory such as cost, speed, storage, and management.
7. Economies of scale. Most people have some sense that the cost per unit goes down with larger scale. Economists explain the concepts using the average cost and marginal costs. In the supply chain, the marginal costs and average costs mostly take specific forms. In this chapter, we will discuss economies of scale directly related to the applications in the supply chain, such as in production, transportation, and storage.
8. Economies of sharing and economies of scope. Sharing can reduce the cost. The supplier can optimize the product assortment for the customer to share. Products and services can share the resources and materials. Economies of scope design a flexible supply chain so that it can supply a variety of products and services.
9. Uncertainty: its cost and reduction. Uncertainty incurs cost and leads to risks. We will discuss the measure of the uncertainty, classify the level of uncertainty in different types of supply-chain activities, and quantify the cost of uncertainty in supply chain. We will then discuss the importance of isolating uncertainty from systematic variations, such as seasonality. We will finally present strategies to reduce uncertainty in the supply chain.
10. Economies of speed and network design. Time is money. We will present the ways to convert speed into monetary measures. We will then discuss the speed and cost of different network designs.
11. Quantitative incentive contracts. We will introduce the quantitative incentive models in an inventory setting.
12. Summary. We will conclude the book by revisiting the strategies applied to the food supply chain and the healthcare supply chain. We will summarize the successes and

challenges outlined in the introduction, with references with links to the technical contents and what to do going forward to link the supply chain to human needs.

Chapters 2, 3, 4, 6, 7, and 8 are quite independent. Chapter 7 is related to Chapter 4. Chapter 9 depends on Chapters 6, 7, and 8. Chapter 10 depends on Chapters 6, 7, 8, and 9. Chapter 11 depends on Chapters 3, 7, and 9.

The case study "Barilla (A)" and an article "Aligning Incentives in the Supply Chain," both from the *Harvard Business School Publishing* (HBSP), are used throughout the book to provide insights into the issues and link the models and strategies to the applications. Readers can purchase these to supplement the contents in this book. References are at the end of Chapter 3.

9.2 The content in the chapters

In the introduction chapter, almost half of the coverage is to draw reader's attention to the issues about human needs and to isolated situations when the financial-driven supply chain does not serve certain human needs well. The reason is that most supply-chain books focus *only* on the financial-driven supply chain. However, we should not dismiss that it is the power of the markets and the financial incentives that helped the supply chain to achieve abundance. We cannot dismiss the goose that laid the golden eggs. Most of the coverages in the chapters will be on the successful strategies for cost reduction and efficiency improvement in the market. In each chapter, we will point out some examples of financial strategies that may not align with human needs.

ACKNOWLEDGEMENTS

I would like to acknowledge hundreds of students who took this class. Together we all learned a lot. I would also like to thank our colleagues in Georgia Tech for many helps. I would specially like to thank Dr. Paul M. Griffin, who first created the Supply Chain Economics course and later helped me with the manuscript.

REVIEW QUESTIONS

1. Please describe a supply chain of something you use frequently. Include the materials it uses, the common marketing channels, and how they are disposed of.
2. Why do supply chains often involve partners?
3. What is the objective most supply-chain engineers and managers try to optimize, and why?
4. Why is logistics important in the supply chain?
5. What is the focus of microeconomics related to supply chains?
6. What are the measures used by economists? Hint: not a financial term!
7. What recent event changed the non-financial measures in economics?
8. What is the role of economics in the supply chain for humans?
9. Please read a few mission statements of the corporation you are most interested in and sew how their mission statements related to human needs.
10. What does triple bottom line objective include? How are they related to human needs?
11. Please give an example of something that is scarce in the United States today.
12. Please give an example where there is significant room for improvement of the supply chain for human needs, even when their financial performance is strong.

2. COLLABORATION IN SUPPLY CHAIN

CSCMP's definition of "supply chain" emphasizes, "Importantly, supply chain also includes coordination and collaboration with channel partners, which can be suppliers, intermediaries, third-party service providers, and customers." All supply-chain professionals know the importance of collaboration. They can often identify the potential synergy among the partners to reduce cost, improve efficiency, or expand the market. However, their proposals often get rejected by their own management or by their partners. This is because suppliers and clients may have different objectives, possess different information, and respond to different incentives. Therefore, to realize the synergy is challenging. This chapter discusses the potential challenges in collaboration between suppliers and clients. We will refer to a case study Barilla SpA (A) [3] and an article Aligning Incentives in Supply Chain [4] from *Harvard Business School Publishing*. The reader should purchase the case and the article to supplement the concepts. The theoretical discussions are in the Principal-Agent model in the next chapter.

1. BROAD PERSPECTIVES IN COLLABORATION AND COMPETITION

Every supply-chain professional knows that the collaboration and coordination among supply-chain partners are important. Since the supplier and client share the total profits gained through their channel, the partners also share contentions.

In the "The Five Competitive Forces that Shape Strategy [1]," Michael Porter laid out the five market forces that every firm in the supply chain faces, shown in Figure 2.1. The forces from the suppliers and buyers, who we call clients in this book, are relevant in this chapter. The forces from the suppliers and clients are not competition. However, they may be an impediment to a higher level of collaboration. The other three forces will be discussed in Chapter 4, Competition.

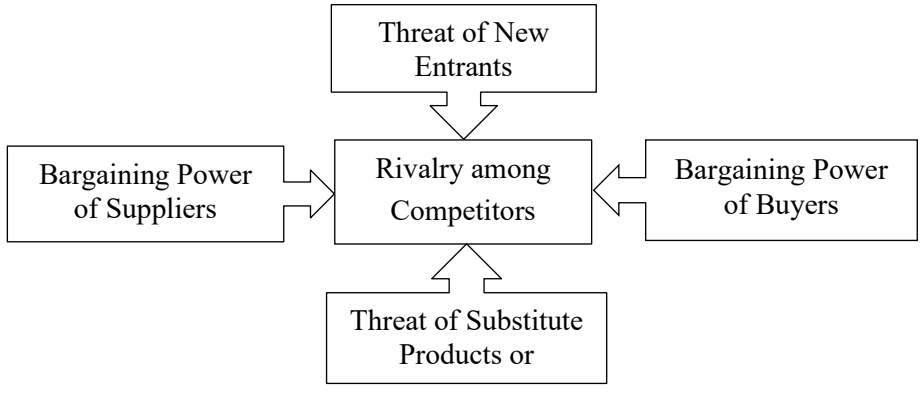


Figure 2.1. Porter's five forces in competition.

1.1 Collaboration challenges with the suppliers

A firm in a supply chain must collaborate with its suppliers to ensure low cost, efficient and reliable supplies, with acceptable lead times. As a consumer, you deal with many suppliers. Imagine you need to get a car. There are many brands, styles, options, and channels to choose from. Your objective would be to find the best supplier with best car, minimum price, and with best shopping experience. To achieve these objectives, you would have to collaborate with the potential supplier. Some suppliers have more bargaining power with you than others. What is the best for you may not be the best for your suppliers? For example, your lowest price may lead to least profit for the supplier. This is true for almost any relationship to a firm's suppliers, either for products or service.

Consider OEM in Figure 1.3 in Chapter 1. There are many potential SCR suppliers with different quality, service, price, reputation, market share, synergy, and financial strength. Each supplier will use their strength to serve the OEM but at the same time, gain bargaining power to get the most from the OEM. The number of suppliers, the availability of substitute suppliers, and the cost of switching suppliers should all be considered in the selection and coordination. For the selected suppliers, the OEM must work carefully to define the liabilities, prices, ownership, terms, and other factors, and monitor during the execution to ensure good quality in products and services.

1.2 Collaboration challenges with the clients

A supply-chain firm must also coordinate with the buyers or the clients. As a student, you may have worked as work-study student, intern, or teaching assistant. In this case, you can be considered the service provider, and your employer is your client. You need to collaborate with the potential employer before the hiring, and collaborate with the employer after you are hired. However, your objectives and concerns may not always be aligned with your employer. The most obvious is that you want the most compensation. However, the higher compensation will cut into your employer's profit.

In the calculator example, the OEM can sell among many potential channels: department stores, specialty stores, campus stores, online retailers, or in-house e-commerce with the objective to maximize profit. Each channel may have multiple potential clients, and each would be willing to pay different prices, sell certain quantities, and with certain service requirements. Each client has different market reach, stability, and capability. They will use their strengths to bargain and gain financial benefits from the OEM. Therefore, the number of potential clients, their strengths, switching cost, profit margin, and synergy are all important in the collaboration strategies.

1.3 Coopetition

Competitors sometimes collaborate. Many retailers such as CVS, Publix, Walmart, and Costco offer their private label products, such as milk, along with the national brands. The retailers work hard to ensure that the national brand is still happy selling their products through their chain, because the national brands increase the store's bottom line. The national brand will also want to ensure the chain stores remain vibrant and able to help them.

Netflix supplies stream video services. It uses Amazon Web Services. Amazon also offers its own video services. Therefore, they compete. However, they must also cooperate. Since Netflix is Amazon's customer, Amazon wants Netflix's to expand to need more services. This is not a zero-

sum game. Together, they compete with DVDs and other channels to grow the market together. Google and Apple compete furiously on smart phones and tablets. However, they collaborate on search engines, maps, and other matters. Competitors sometimes can find common interest from collaboration.

Mercury, a division of Ford, cooperated with Nissan to produce minivans. Neither firm could supply the minivan successfully on its own. Trade organizations get together to share information on markets, form coalitions to work with common suppliers, and expand the client base for all members.

Brandenburger-Nalebuff proposed a value net, shown in Figure 2.2, which depicts the concept of coopetition [2]. A firm may have competitors that share the same suppliers and customers but no other common interest. The competition is a zero-sum game. The "Coopetitor" is different from a competitor in the sense that they share some common values, as shown in the examples above.

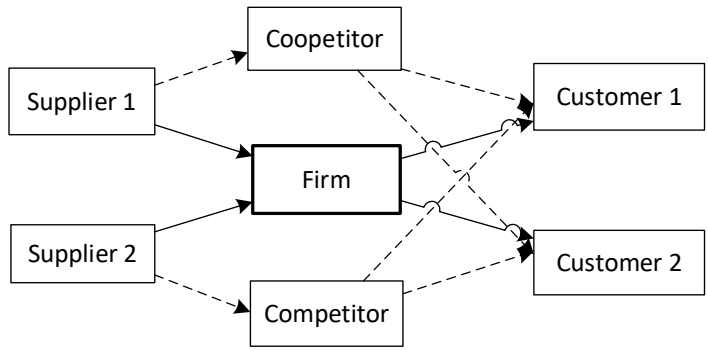


Figure 2.2. The relationships in coopetition.

In United States, there are five major freight rail companies. Each company concentrates in certain regions in the country with some shared territories. They compete in the shared territories. However, the origin and destination of a large portion of freight demands are covered by regions owned by different companies. Therefore, they have to collaborate. Otherwise, the customer may use trucking services instead. In the logistics section, you will see that these are also the most suitable demand for rail.

Coopetition can go too far to become collusion. The U.S. government filed an anti-trust lawsuit against Google in July 2020, in part related to Google's exclusive business deals with Apple.

1.4 Factors that impact the collaborations

There are concrete contentions between the partners. Here are a few examples.

1. The profit of the seller increases with the selling price, while the profit of the client decreases with selling price.
2. The firms would not reveal information important for collaboration, such as the internal costs, external costs, the business model, future plans, their partners, and matters.

3. Most suppliers in a supply chain have multiple clients. A supplier's financial objective is to maximize profit among all the clients in both the short and long term. The financial rewards from different clients are different. Most clients have multiple suppliers. Their financial interests mirror of those of the suppliers.
4. The relationships between a supplier and a client are defined through formal or informal contracts. However, these relationships are hidden from the others.
5. There is always uncertainty involved in supply chain. The supply of products and services takes time—often days, weeks, or months. Therefore, the investment in resources and materials expose the firms to risks. Which partner should take how much risk is a major contention between the channel partners?

These challenges arise from objective and information asymmetry.

2. THE DRY PASTA SUPPLY CHAIN

We will illustrate the challenges in the collaborations between a pasta supply chain with a producer, a food distributor, and supermarkets. The reader can read Barilla SpA (A) to learn the details, including each partner's objectives, incentives, information, personal interest, and other insights related to collaboration. You can purchase the case from Harvard Business School Publishing directly. If you are taking a class using this book, you can ask the instructor to set up the purchase to enjoy the discounted rate.

2.1 Dry pasta production and distribution

The dry pasta suppliers produce and sell the pasta to distributors. The distributors sell pasta, along with other products, to supermarkets, which sell products from this distributor and other suppliers to consumers.

2.1.1 Production process

A pasta producer produces many types of dry pastas of different ingredients and shapes. The mass-produced dry pastas are produced on large production lines. You can find videos of large-scale pasta production lines on the internet. A large production line can produce many truckloads of pastas per day! The process in a production line is shown in Figure 2.3.

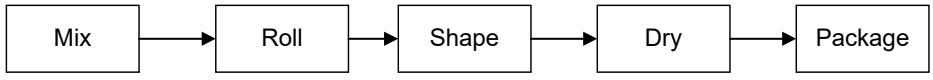


Figure 2.3. The processes in a pasta production line.

2.1.2 Pasta distribution

After production, the dry pasta will be stored in factory distribution centers. The distributors will order pasta from the producer periodically. The distributor will then sell the pasta, along with products from other suppliers the distributor carries, to the supermarkets, who sell pasta along with many other products to consumers. Supermarkets may also carry pasta from other suppliers.

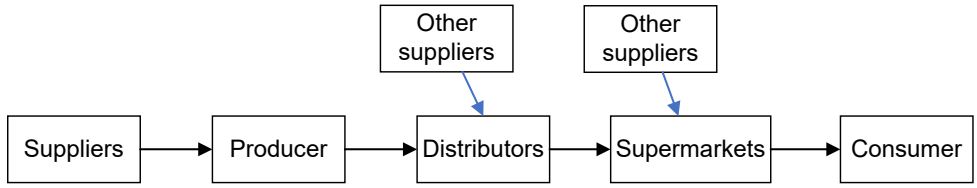


Figure 2.4. Dry pasta supply chain.

2.2 Producer marketing

The staple food-supply chain is a very competitive market with a large and steady demand. There are many producers competing to earn profits. Dry pasta is a low-margin, high-volume product, bulky, and with a long shelf-life. To compete, the large food producers often use several strategies:

1. Merchandizing. They send their salespeople to help the retailers in product display and promotion, and set ordering strategies. For example, Proctor and Gamble, a major supplier of many consumer products, invests heavily in merchandizing. The result is obvious.
2. Promotions. They use various ways of advertising using select contents through select media through different campaigns.
3. Periodic discounts. The producer offers discounts, coupons, or rebates on certain types of products periodically based on their supplies, production, inventory position, or market situations.
4. Quantity discount. This could be buy-one-get-one free, same price for larger quantities, or a discount on a caseload, truckload, and so on.
5. Transportation discount. This might be free shipping.

2.3 The performance indicators

Performance can be examined by the order patterns, inventory levels, and stockouts. High-performing supply chains would show steady orders, low inventories, and minimum stockouts. The orders Barilla received from each distributor is highly variable: one week a lot, next week very little. One of the useful measures of variability is the coefficient of variation, CV, which is defined as the ratio of the standard deviation and the mean μ, or

$$CV = \frac{\sigma}{\mu}$$

When CV is greater than 1.33, the variability is considered high. When CV is between 0.75 and 1.33, the variability level is considered moderate. The average weekly orders of all pasta types of the producer received from one of the distributors is in the moderate level. The individual pastas can be in the high variability range because the variability decreases when pooled among many different pastas or different stock-keeping units (SKUs).

One may think the high variability was caused by the high order variability from the supermarkets. However, the coefficient of variation from supermarkets shows $CV \ll 0.75$. In the case study, the coefficient of variation is approximately 0.2.

One may think the reason for high order variability in what the producer received from the distributor could be that the distributor does not carry sufficient inventories to buffer against uncertainties. The case shows that the distributor carries about three weeks of inventory, which is not low for this type of staple product. Even with so much inventory, the distributor still suffered from significant stockouts!

2.4 The problems with high order variability

Highly variable demands from distributors cause major challenges to the supplier's production system. In food, carpets, paper, and many other consumer products, the setup costs due to changeovers between different types can take a long time. The product during the transition may not meet the quality standard and therefore generate waste. The setup costs depend on the sequence of successive products. Therefore, the production planner would optimize the sequence and lot sizes to reduce the setup costs and inventory-related costs. However, lumpy large orders can deplete the inventory of a specific pasta very quickly. The production manager may have to interrupt their production plan and run the emergency batch to avoid shortages. The change would lead to higher setup costs, undesirable batch sizes, and with possible stockouts of other products.

2.5 A proposal and its reception

Since the actual demand from the supermarkets to the distributor has low variability, opportunities exist for the producer, the distributor, and the supermarket to collaborate to reduce the variability of the orders from the supermarkets to the distributor, to reduce the amount of inventory and stockouts at the distributor, and to provide better service to the supermarket. The cost savings resulting from reduced production setup costs, inventory costs, and stockout can lead to lower prices and more sales. The reduced cost can allow lower prices at the supermarkets, which can increase the demand and profit for all partners to share. It should be a win-win-win for all three partners. This is similar to the Just-in-Time (JIT) production system from Toyota.

However, the idea was not well received by the distributor. The distributors did not want to lose the ability to enjoy the discount under their control. Even the salesforce resisted the proposal. They call the proposal infeasible and dangerous.

2.6 Discussions

There are three partners in this chain: the supplier, the distributor, and the supermarkets. The excess cost in production, inventory, and stockouts can lead to lower profits, loss of sales, loss of market share, and can tarnish brand image for all. A streamlined supply chain can benefit all parties. However, the apparent good proposal was not well received by some branches inside the producer and by the partners. What can be done to realize the improvement potential?

3. METHODS TO IMPROVE COLLABORATION

Supply-chain professionals may find opportunities in cost reduction or efficiency improvement across partners. They can also find opportunities in increasing profit through pricing or market share. However, to convince the management within the organization may be challenging,

because different branches or even the management may have different incentives. To convince the partners is even harder without trust. We will outline some methods useful to develop proposals others can accept.

People and organizations respond to incentives. The incentive alignment is important for partners in the supply channel to coordinate well. We will use a *Harvard Business Review* article [4] as reference. Similar to the Barilla Case, the reader should purchase the paper to get a better understanding of the concepts.

3.1 Identify opportunities

First, opportunities must exist through collaboration, such as reduced total cost among all partners, faster service to the end-consumer across the partners, higher quality for better prices by the client, or lower prices at the end of the chain for increased market share. In the dry pasta case, smoother demand from the distributor to the producer can reduce the production cost. The cost savings can be used to lower the prices to the distributors and to the supermarkets to increase the market share. The reduced inventory at the distributor can reduce the cost to the distributor, who can pass on the savings to the supermarkets. The streamlined supply chain can also possibly reduce the lead times, which can lead to reduced risks, reduced stockouts, and better services. Therefore, synergy exists among the partners.

3.2 Challenges in incentive alignment

The major challenges every player in the supply must face include the following

1. Objective misalignments between channel partners. Each player in a supply-chain channel has its own objectives. One of the objectives for all parties is to maximize their own profit. The producer's profit margin increases with the selling prices to the distributor or larger market share through lower prices at the supermarket. The distributor's profits increase with lower purchasing prices for this low-margin product. They want to maximize the discount they can enjoy, but have no incentive to discount the selling prices to the supermarkets. The distributor also wants to maximize profit among products from all suppliers, not just the pasta supplier. The supermarkets have even more suppliers to draw profit from. In general, the objectives of the partners share commonalities but also have differences. We call the dominating partner the Principal and the dominated partner the Agent. The Venn diagram of their objectives is shown in Figure 2.5.

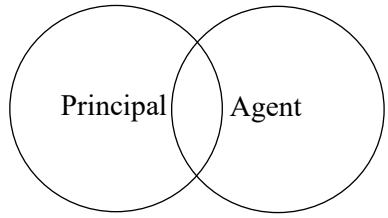


Figure 2.5. The objectives of Barilla and the distributor.

2. Objective misalignment between branches in a firm. Different branches in the same firm may have different incentives and objectives. In Barilla's case, the production manager wants to minimize production and inventory-related cost. The logistics manager wants to reduce transportation cost, and the salesperson wants to maximize the bonus measured against campaign goals. The salesperson would also like to increase the SKUs to increase the sales. However, the SKU proliferation can cannibalize the sales of other SKUs and lead to higher cost in production and the supply chain.
3. Information asymmetry. Each partner or branch within a partner in the channel has access to information related to its suppliers, its clients, prices, costs, production plans, and process information. Some information is proprietary; others can be public due to regulations or business needs; and still other firms are often treated as proprietary but in reality, can benefit the strategic partners. However, careful analysis is needed to ensure that the sharing will not compromise a firm's own interest. For example, if the supermarkets share their aggregate daily sales of different pastas with the distributor and producer, the distributor and producer can use the information to plan their inventory, order, and production. The supplier and the distributor may offer discount to the supermarkets for sharing the information. On the other hand, too much information—such as the price and all items a customer bought—should not be shared, because such information can compromise the supermarket's own interests.
4. Risks. Every firm or person faces risks, because
    a. The future is uncertain, demand is uncertain, and they vary with seasons and events.
    b. All processes are subject to interruptions and uncertainty.
    c. Production, transportation, order processing, and delivery take time.
    d. Order process takes time. After an order is received, the supplier must plan when to fill the order and spend the time to fill the order.
5. Human behavior. Because of misalignment, information asymmetry, risks, and uncertainty, people take actions toward their objectives. Some information they possess may not be revealed to the others, and some of their actions are not observable by the others.

To mitigate the potential damages stemming from the risks caused by these factors, firms can keep safety stocks, plan for excess capacity, or plan for extra time to hedge against the uncertainty. However, these measures cost resources. Each firm would hope their channel partners would invest to hedge the risks instead of themselves.

3.3 Identify the differences

It is very important to identify the differences objectively. It is difficult to be in other's shoes.

3.3.1 Identify objective misalignment

First, the partner that intends to improve collaboration must identify the similarities and differences between the partner's objectives. In the challenges in the Barilla case, we have discussed that all parties wanted to increase profit. However, their objectives share similarities and have differences. One specific example in the case study is that Barilla discounts its prices to the distributor with the objective to lower the prices at the supermarket, to increase market share. However, the objective of the distributor is to maximize the discount to increase the margin between buying and selling. Worse yet, the periodic and volumetric discount caused high-order variability from the distributor with large orders at discounted prices or no orders at other times.

3.3.2 Identify information asymmetry

The information privacy that leads to asymmetry may be necessary for the agent to compete in its own market. In Barilla's case, Barilla would not share how much they pay for their ingredients, the cost of production, or wages of salespeople because that can compromise Barilla's interests. The distributor would not share their detailed selling prices to the supermarkets and their relationships with other suppliers. On the other hand, the daily stock levels and the sales from the distributor to the supermarkets may not hurt any partner.

Supply-chain information sharing is difficult because of high volume and difficulty to integrate into the partner's software systems. The sharing may require significant investment and expertise.

3.3.3 Identify hidden actions

Each firm has many people who do many things every day based on the incentives provided. Some actions are not observable by their partner or even within an organization. A production manager may run a batch longer to reduce setup cost even when there are emergency needs for something else. A salesperson pushes a large sale at the end of month to ensure a bonus, but that leads to a huge order that cause stress in manufacturing and logistics. The actions by the people in the partner organization are normally not observable—for example, a distributor pushes a competing brand due to its promotions.

3.4 Possible solution strategies

Several solution strategies are discussed in this section.

3.4.1 Contract with clearly defined responsibilities

The first possible strategy is to develop a contract that specifies responsibilities, ownership, and payment structure clearly based on observable information and enforceable actions. Cisco's 2.5-billion-dollar inventory write-off during the internet bubble did not satisfy any of these. Its contract did not clearly stipulate the responsibilities of inventories. In addition, the strong push for fast delivery in its long supply chain in an expanding market incentivized overstock of supplies along the chain, especially near the point of consumption.

3.4.2 Information-sharing strategy

A second possible strategy is to identify the information critical to align the incentive and the partners cannot be hurt if shared. In the Barilla case, the distributor's weekly sales of all the SKUs to the supermarkets is very useful for Barilla to plan for its production and delivery, reduce production cost, reduce delivery cost, and avoid stockouts for everyone. However, a distributor's competitors can use this information to take advantage of the distributor. However, as a supplier, Barilla cannot use the distributor's sales data to take advantage of the distributor. If Barilla can gain the trust from the distributor, and sign the non-disclosure of information, and can demonstrate that the sale's data from distributor can benefit the distributor with quantitative analysis, such as lower average prices, lower inventories, less stockouts, the distributor may be more amenable to the JIT distribution. This is achievable because of the total cost reduction among the partners and increased sales through reduction of stockouts. If the cost reduction is shared fairly, what is the reason the distributor would refuse?

Campbell's case study outlined in the HBR's incentive paper discussed the benefits of information sharing. In the 1980s, Campbell's promoted sales by offering discounts to distributors. It hoped that the discount would be passed to the retailers and increase the total sales. However, the distributor's margin is only 1–3%. The distributor can increase the profit in higher percentages if they can buy during the discount. As a result, whenever there is discount, the distributor takes the opportunity to place bulk orders and stock up the inventories. They then enjoy a higher profit margin without passing the discount to retailers. This leads to demand fluctuation without demand increase. Campbell identified the problem observing distributor's pricing information on both buying and selling. They then identified the source of information asymmetry and misaligned incentives, and proposed flat low prices in exchange for the information on the scan at the sales. The scans are frequent and observable action, and the data can be shared without hurting the distributor. Scan data sharing can be a challenge if shared manually due to the complexity of SKU, ID, time stamp, and so on. Some form of automation is often necessary. You can find Campbell's results in the next section.

3.4.3 Trust-based strategies

Another strategy is to build trust. Sometimes firms maximize interest based on the short-term objectives. Since things can change quickly, the partnerships are based only on financial terms. However, in the long run, the lack of trust does not help to deter the adverse selection and moral hazard. If the firms take a longer-term view to build the trust, the partners may establish more trust. For example, giving Campbell's access to distributors pricing in both buying and selling requires a lot of trust. Therefore, a long-term contract with frequent information exchange can provide the best outcomes. Another way to acquire trust is through intermediaries. If there is a third party that both the principal and agent can trust, they can achieve trust to deter adverse selection and moral hazard indirectly because the intermediary's reputation is on the line. The consulting firms, supply-chain service providers, and even mutual friends can be such intermediaries.

3.4.4 Incentive strategy

Incentives can be built into the contracts to align the objectives and to reduce information asymmetry. One of the objective misalignments between Barilla and the distributor was that the discount intended for lower prices at the supermarket became the increased margin for the

distributor. Barilla can change their pricing strategies to incentivize the distributor to lower their prices at the supermarket while the distributor still enjoys the same profit or more. The incentive can be linked to sales prices or volumes at the supermarket or anything observable by Barilla. You can find a quantitative example in the HBR article. It will also require certain information transparency among the producer, the distributor, and the supermarkets.

3.5 Invest in information-sharing

The quantity and prices of each SKU at the clients are very useful information for the supplier. However, the sharing of such information securely among many SKUs frequently and using them for planning and execution are complicated and require significant resources. The Electronic Data Interchange was developed to allow data interchange among different supply-chain applications to share data automatically. However, not every firm has that capability.

Sam Walton, the founder of Walmart, was the first to recognize that the point of sale (POS) data contains a wealth of information useful to improve the entire supply chain. At the cash register, specific stock keeping units (SKUs) are checked out by known quantity at a known location at a known time. These are the key attributes of the demand and inventory status at the store! Imagine how useful these data can be for Barilla and the distributor! They can use the data to plan for their production, replenishment, order fulfillment, and transportation.

POS also provides the price, the other items a consumer buys at the same time, credit card information, and so on. These are extremely important to learn about the affinities between SKUs, shoppers, and membership. They are useful for the retailers in product assortment, merchandizing, advertisement, and many other functions. However, the supermarket may not want to distribute such information to Barilla or the distributor.

4. EXAMPLES

We will use a few examples to illustrate the issues and solutions.

4.1 Waste management in conference centers

A senior design project worked with a large conference center. The conference center occupies a show space of over one-million square feet, and entertains over a million visitors every year. During setup, break down and the conference, tons of packaging materials, construction and demolition materials, food waste, and so on, are disposed of. In 2018, only 87% of the waste was sent to landfill. The conference centers would like to increase the percentage of waste to be sorted and diverted from the landfill by designing new contracts with the conference hosts.

Many parties are involved in setting up and tearing down events. Conference hosts, such as MHI, contract with the conference center to hold their events. The exhibitors such as Dematic or Manhattan Associates contract with the conference host to set up booths to display their products or services. The exhibitors would hire contractors to set up and tear down the booths. The conference center would also hire general contractors to set up the common facilities such as aisles or power supply before the conference, and clean and maintain the conference center space.

The new contracts are for conference hosts and waste haul services. The conference center does not have direct contracts with many subcontractors, shown in dashed boxes, who impact the waste diversion. The detailed relationships are shown in Figure 2.6.

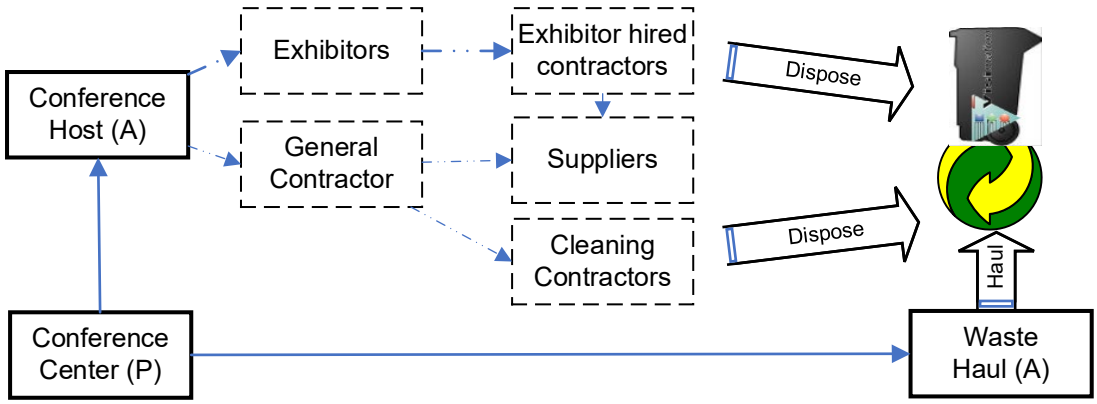


Figure 2.6. The contractual relationships of the players that generate and sort the waste.

Table 2.1. The roles and objectives of different players in waste management in conference centers.

Player	Role	Objectives	Actions and Visibility
Conference center	Landlord	Minimize landfill and cost, maximize diversion	See transportation charges for disposals in waste diverted and penalty for contamination
Conference hosts	Organize conference	Quality conference, minimum cost	Contracts with exhibitors and general contractors, observe the conference
Exhibitors	Hire contractor to set up booths	Best-looking booth with minimum cost	Incentive to use single-use carpet, furniture
General contractor	Set up and breakdown the common areas	Minimize time spent and cost	Dispose of large amount of single-use carpet for aisles
Exhibitor-hired contractors	Set up and breakdown booths	Minimize time spent and cost	Incentive to maximize single-use materials
Suppliers	Supply materials to contractors	Maximum profit	Incentive to select single-use materials

Player	Role	Objectives	Actions and Visibility
Cleaning contractors	Clean up the facilities	Maximize billable hours	Sorting is not part of the contract, and even if it is included, is difficult to observe
Waste haul	Haul waste or divert containers to their destinations	Minimize effort to get paid by the contract	Know if the load is waste or diverted, and if the diverted load is contaminated and charge conference center

Currently, the waste haul charges the conference center $400 per trash haul. The waste haul charges less for the diverted materials because they can sell to respective markets. The prices of diverted materials fluctuate in a local market. In Georgia, it has been approximately $200 for construction and demolition, $185 for plastic sheeting, $43 for cardboard, and $0 for furniture.

To get a better understanding, readers are encouraged to do a role play to see what should be included in the contracts to the conference hosts and waste haul to incentivize higher diversion rates. The diversion rate can be estimated and verified later by the waste haul.

Most contractors nowadays are concerned about sustainability. Sorting may not add much extra work during certain steps during the setup and teardown. Once mixed into a dumpster, it is much more difficult to sort out.

Since the objective of this project is about sustainability, and all corporations strive to achieve sustainability measures, what other incentives, financial or otherwise, can you use to increase the diversion rate?

4.2 Barilla's case study

Please read the case study and the incentive alignment article carefully, and then fill out table 2.2.

Table 2.2. The objectives and other issues of different parties in the dry pasta supply chain.

	Producer			**Distributor**
Issue \ Role	Production Manager	SC manager	Salespeople	DC
Objective				
Inventories				
Private information (Some do not have to be private)				

	Producer			Distributor
Actions hidden				
Information gathering				
SKUs				

The reader is encouraged to play different roles again to get a better understanding why JITD was not well received by some, and what can be done to align everyone's objectives.

4.3 Other examples in the HBR Article

1. Campbell's Soup example, details in another HBS case, somewhat similar to Barilla case with initial results of changes through level pricing, information sharing, and JIT-style deliveries.
2. A bread manufacturer's drivers were allocated a specified amount of shelf-space, and paid a commission based on the sales generated at each store served. Because they were not penalized for the stale inventory wasted, they filled up the shelves even on the days when completing brands were offering substantial sales.
3. Tweeter electronics store salespeople push other's products, such as Tweeter.
4. Norwalk made special furniture to try to sell via retailers. They found the stores often pushed for their own products instead of Norwalk's based on what was in the inventory.
5. SoundScan collects retailers' sales data and sold to manufacturers and retailers. There are trendsetter stores and trend-follower stores. The trendsetters provide good forecast for the industry. However, they have little to gain to join the group. The trend followers and manufacturers can use the information to provide more accurate forecasting, and therefore have more to gain. However, if trend-setters do not participate, the information collected will become a rear-view mirror, not useful for the trend-followers and manufacturers.
6. Supply-chain partners will try to hide their processes, costs, plans, and so on, from the others to avoid being squeezed of the profit margin.
7. On page 7 of the HBR article, there is a quantitative example of a newspaper publisher and newsvendor. The print cost is 45 cents per copy, wholesale price is 80 cents, and retail is $1. The salvage value is 0. The demand follows uniform distribution $U(100, 200)$. The vendor can make 20 cents for each copy sold but loses 80 cents for each copy not sold. If the publisher and the vendor optimize their own utilities in a traditional contract, the newsvendor model will lead to the result that the vendor would order 120 copies. The vendor will make $22, the publisher $42. If the publisher offers a revenue-sharing contract, the publisher lowers the wholesale price to 45 cents (at cost); the vendor returns 35 cents back to the publisher for each copy sold (winds up paying the wholesale only for the copy sold); the optimum order quantity is 131; publisher profit increases 5.17% while the vendor profit goes up 4.9%. The publisher can also pay the vendor 60 cents for

a copy unsold. In this case, the order quantity is 150 copies. The publisher profit increases 7.1% and the vendor profit 13.6%.

4.4 Various forms of supply-chain contracts

There are many supply-chain contract types. The HBR article listed two numerical examples in which the supplier designed a contract.

1. Revenue-share contracts. The supplier lowers wholesale price and the client shares the revenue with the supplier for each unit sold by a predetermined amount in the contract.
2. Buy-back contract. The HBR article called it a mark-down contract. The supplier would buy back the unsold units from the clients at an amount predetermined in the contract.

There are situations in which the client would design a contract in a similar vein:

3. Pay-back contract. In this case, the client would want to ensure sufficient supply, and would pay the supplier for the units not purchased from the supplier at a value higher than the salvage value realizable by the supplier.
4. Cost-share contract. The client would pay some supply cost with the supplier to increase the supply quantity. In return, it enjoys some wholesale discount.

The other forms of supply chain contracts can be:

5. The client orders a percentage of production capacity without specific quantities on individual SKUs. This will give the client some flexibility in the product assortment while maintain the supplier's revenues.
6. Flexible quantity contract. This is to allow fluctuations within a range, possibly at different prices. This will allow both parties to limit the risks.
7. Portfolio contract. These design multiple contracts to allow flexibility to hedge against risks.
8. Performance-based contract. We will discuss a simple two-level contract in the next chapter.

In Chapter 11, we will develop models and perform detailed analysis of the first four contracts, after we build up the background in the following chapters.

5. COLLABORATION AND HUMAN NEEDS

Most firms today add objectives in social responsibility and environmental stewardship. These objectives are closely tied to human needs. These objectives may not have short-term financial payoffs. However, they can have significant long-term financial implications. Overlooking such objectives can also lead to costs for the firms. The supply-chain professionals can incorporate these into their supply-chain strategies.

The example of waste management in conference centers illustrates the effective use of such strategies. Their incentives using awards do not incur much cost but can lead to significant gains in reputation, in public relations, and bottom lines in the future for the conference hosts.

When the financial utility is not aligned with human needs in collaboration between the supply-chain partners, the successful collaboration may not lead to improvement in satisfying human needs. Food deserts are an example.

As the society becomes more affluent, especially when income inequality grows, these two utilities may differ more and more. For example, in the pasta supply chain, pushing for over-consumption of tasty but unhealthy pastas can lead to financial gains but hurt the human needs. The food desert and healthcare supply chains are two examples in which the financial utility in the collaboration between the supply-chain partners does not align with the human needs utility. If firms want to serve human needs, they must find ways to internalize some human needs measure into their financial systems.

3. PRINCIPAL-AGENT MODEL

Firms in a supply chain must collaborate with its suppliers and clients. To collaborate well, a firm must find the similarities and differences in objectives and develop strategies to incentivize the sharing of useful information and to guide the hidden actions to benefit the partners. In the Barilla's case, the misalignment of the goals and incentives and lack of sharing of useful information and wrong incentive guided actions among Barilla, distributors, and supermarkets led to wild order fluctuations, high inventories, and high stockouts.

The principal-agent model in agency theory and behavior economics has been developed to understand the information, information asymmetry, and behaviors that occur between the collaborating partners, and to develop strategies to use incentives to line up the objectives between the partners. Behavior economics has gained more and more attention in recent years. Psychologist and economist Daniel Kahneman and later Richard Thaler have won Nobel Prizes in 2002 and 2017 in economics, respectively, for their works. The understanding of human or business behavior is extremely important in supply-chain design.

1. INFORMATION AND BEHAVIOR

Many say that the society has entered the era of knowledge economy. The information is the foundation of the knowledge economy. Many supply chain firms find way to find data, analyze the data to predictive market information and customer behavior. There are different types of information and behavior relevant to supply chain.

1.1 Public information

Public information is accessible by everyone. It can include information on websites, social media pages, weather forecasts, published prices, government publications such as from the Bureau of Labor Statistics, CDC, or Centers for Medicare & Medicaid Services, or SEC's 10K annual reports of the publicly traded companies.

Data available publicly may not always be in a form convenient for analysis. Skills from web-scraping to data-mining can be useful to convert the raw data into organized data for analysis. Firms with these skills would have an advantage over the ones without. For example, the supermarket prices in Barilla's case are public for everyone to see in the store, including the distributor and Barilla. However, they are difficult to collect and update into an organized form for analysis. Similarly, prices online are public information. But it is difficult to get them organized for analysis. This difficulty is an opportunity for entrepreneurs to develop Apps to help potential users.

1.2 Private information

Some information must be kept private or proprietary to protect the owner's self-interest. In competition, revealing the planned supply quantity and supply cost can impact a firm's profit! Other private information can be detailed cost structure, compensations, and research and investment plans. For example, Barilla would not share the detailed cost and compensation information with the distributors or the supermarket, because the latter can use the information to

take advantage of Barilla. Barilla would also keep such information from competitors to maintain its profit margins.

A firm's specific relationships with each supplier and client—such as discount, cost share, and revenue share—are important to the firm's mission and are normally private. For example, the pasta distributors in Barilla's case would not share their pricing and discount deals with Barilla.

Some student may consider personal health situation, or the commitment to activities that may potentially interfere with the effort available to contribute to senior design, private.

The private information and the strategies to deal with it are the most important part of collaboration and in the principal-agent model.

1.3 Laws on privacy

There are laws or regulations that require the revealing of information, such as K-10 report, enrollment, job placement, and so on.

There are also laws to protect information. Here are a few federal laws for private information protection. There are also state-level laws.

The Health Insurance Portability and Accountability Act of 1996 (HIPAA) protects patient health information. There are drawbacks. If I passed out in class, some may call 911. Maybe it is a chronical situation for which I have medicine in my pocket for quick relief. If others know, they can help me. HIPAA also helps to make the healthcare market opaque.

The Family Educational Rights and Privacy Act (FERPA) protects student privacy. However, the concerned parents who pay the bill can sometimes be left in the dark.

The Children's Online Privacy Protection Act (COPPA) protects the personal identifiable information (PII) for children 12 years or younger.

The Gramm-Leach-Bliley Act (GLBA) protects nonpublic personal information (NPI) such as property records or mortgage information. Banks periodically send notifications to explain the categories of NPI that are being collected and shared, and opt-out instructions. However, there is a loophole that the consumers do not have control with third-party companies affiliated with the bank or insurance companies.

1.4 The cost of information gathering, organizing, and sharing

Each day, a manufacturer generates a lot of information related to the different products that are important to the efficiency of the supply chain. A 3PL deals with thousands of products, hundreds of transporters, and many suppliers and clients. A store serves hundreds of customers who purchase thousands of items. Each item purchased embedded a lot of useful information for the supply chain. Millions of things change status frequently. To collect the detailed and complex information automatically requires resources.

Some professional organizations, such as CSCMP, publish data useful for their members, and sell reports. Many consulting companies collect useful data for sale as part of their business model. Those who use the information frequently, such as some consulting firms, can justify the cost, while an occasional user cannot justify the cost.

Some domain knowledge takes years to acquire and practice, such as the knowledge of a competent supply-chain professional, a doctor, a consultant, or a freight forwarder who knows a

lot about the freight rate, time, customs, and even the people they have to work with to get the goods across borders. Such domain knowledge gives professionals an advantage in their dealings with others.

Supply-chain data are complex and extensive in volume. The sharing of information among different applications in different firms requires infrastructure, expertise, and customization. The cost of sharing can be significant.

1.5 Information asymmetry

When you are buying a car, do you wish to know the true value of the car to the owner? When your doctor recommends a treatment, do you wish to know the margins that the doctor and doctor's office receive? Do you wish to know the contracts between the doctor's office and your insurance company? Your doctor, the doctor's office, and the insurance company do know these things. This is information asymmetry.

When two firms collaborate, there is public information, each firm's private information not to be shared, and the information the firms negotiate to share. The Venn diagram in Figure 3.1 shows the relationships. X and Y denote private information for firms X and Y; O denotes public information; and XO denotes the public information that X has organized and used for analysis and decisions. XYO denotes the public information that both X and Y have organized and used for their decisions. XY denotes the part of the private information that firm X and firm Y negotiate to share. This is the most important part in Principal-Agent model.

In addition to information asymmetry in X and Y, there is also information asymmetry in XO and YO. For example, when a startup needs services from a well-established 3PL, the startup would know much less than the 3PL about public information because it will not have time to collect and to digest the information. On the other hand, the 3PL may have internalized, probably with automated updating, the public information such as weather, road conditions, and so on into its ERP system.

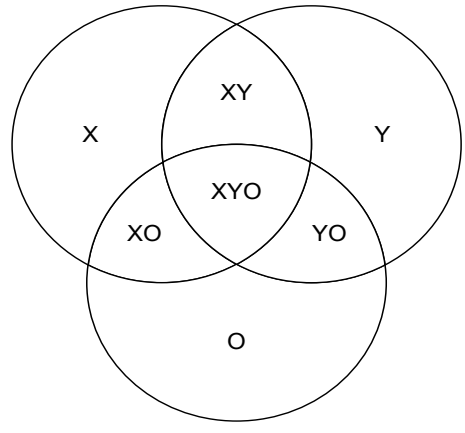


Figure 3.1. The private information between two firms and in relation with public information.

1.6 Information sharing

Some information is easy to share. A job seeker can provide the resume. A small party planner can provide the list of the plan to the caterer.

The supply chain critical information between collaborating partners such as manufacturers, 3PLs, distributors, and stores is high in volume and complexity. Each firm runs its own enterprise system, such as enterprise resource planning systems (ERPs). Each system has its own data structure with its own definitions of terms. To be effective in collaboration, the systems have to share large volumes of complex data frequently, such as each day.

Electronic Data Interchange (EDI) is a set of standards that define different fields and formats that most modern software systems support. HTML allows all browsers to understand where to put what, and in what way, from a text file. EDI allows business data to be understood between different software systems. EDI was initially supported by its own markup language, but today that function is done through XML. When two firms are linked by EDI, the SKU, quantity sold or bought, time, an order, shipping address, and so on, can be exchanged and directly put into their own database system for decision support.

Firms must determine what information to share with which partner. For example, Walmart initially shared certain sales data, such as time of day and items bought together, with the suppliers; and later determined that those are important for its own core competency. The data not shared are private from its collaborative partners.

1.7 Signaling

Signaling refers to the agent's effort to reveal or emphasize certain information to the potential partners. Sending a resume to apply for a position is signaling. Some information on the resume may not be public information. Some people even send somewhat different resumes for different jobs, to signal the different message. Companies may display their awards and certifications, and adopt certain design styles, to send out signals for potential collaborators. Bank headquarters want to show off financial strength. Walmart wanted to show off its thrift culture in its current headquarters, and community focus in the new home office campus.

1.8 Adverse selection

Adverse selection is when a partner selects not to reveal certain private information to the other side. In Barilla's case, Barilla would not reveal its dealings with other independent distributors and supermarket chains to a specific distributor. The distributor would not share its dealings and priorities with different suppliers, such as Barilla. A 3PL would not reveal the prices or return policies with other clients. A firm would not reveal former bankruptcies, criminal record, or debts. For individuals, one may decide not to include unfavorable information in the resume. The most common example used in the economics textbooks is a person buying health or life insurance choosing not to reveal chronic disease. People can also adversely select signals such as different resumes for different jobs prospect.

1.9 Moral hazard

Moral hazard refers to the actions one of the partners is taking in a task that can hurt the other partner without hurting himself/herself. Here are some examples. A supermarket pushes for the

sales of a leading pasta instead of Barilla's. A student team member prepares his/her own exam during team meeting, knowing that the action is not observable by the others. Another example is a fixed-income employee shirking on the job knowing the action cannot be traced down to individuals. A service provider may play favors among clients as long as their own interests are observed. A supplier cuts corners in its products that do not have short-term consequences. Wasting electricity, hot water in the dorm, or food in all-you-can-eat buffets. To drive rental cars hard. Hospitals waste supplies or over-prescribe if they are reimbursed by the insurance. The insurance pays higher charges to hospitals than needed because they can collect the money from premiums. Or unnecessary visits to health services knowing that the insurance will cover all the cost.

2. THE PRINCIPAL-AGENT MODEL

When a firm is to find a partner or find a firm to perform some tasks, it will develop a procedure to find a partner. Those interested may respond. Interactions ensue. After some rounds, an agreement may be reached on the types of services to be rendered, the duration, the outcomes, and the payment scheme, often in the form of a contract.

2.1 Principal and agent

We will look at the simplest case in which there are two players: a principal P, and an agent A. The principal has a task that needs to be executed by someone else. The principal initiates the process to find an agent. The agent will execute the task. In the Barilla case, Barilla was the principal while the distributor was the agent. Barilla was exploring how to change the contract to improve performance of the supply chain.

2.2 Important milestones in contracting process

The principal-agent (P-A) model is depicted through milestones in the contracting process shown in Figure 3.2. The principal's activities are shown above the milestones and the agent below.

It first performs some market study to find the relevant publicly available information. Many potential agents constantly look out for, or expand, their opportunities. They tried to put out the best information as signals to the potential principals. In order to find the best agents, the principal also show its signals as public information. Based on the public information and signals, the principal will design a contract. This can be in the form of an advertisement. In the business world, this is often in the form of call-for-proposals (CFP). The agent, such as a distributor, a supplier, a service provider or a graduate, can engage the principal, review the contract, and accept or reject the contract. If reject, the process stops for the agent. If the agent accepts the contract, the partnership forms. The agent will execute the tasks outlined in the contract. The outcome of the tasks performed depends on the agent actions and Nature, which is outside the agent's control. The actions can be the amount of time and effort that the agent puts in to enhance the outcome, such as more sales. Nature can be many factors such as market situations. A booming market can enhance the outcome, while a pandemic can shrink the outcome. At the end of the contract, the principal and the agent will share the payoffs from the tasks, such as profit from total sales.

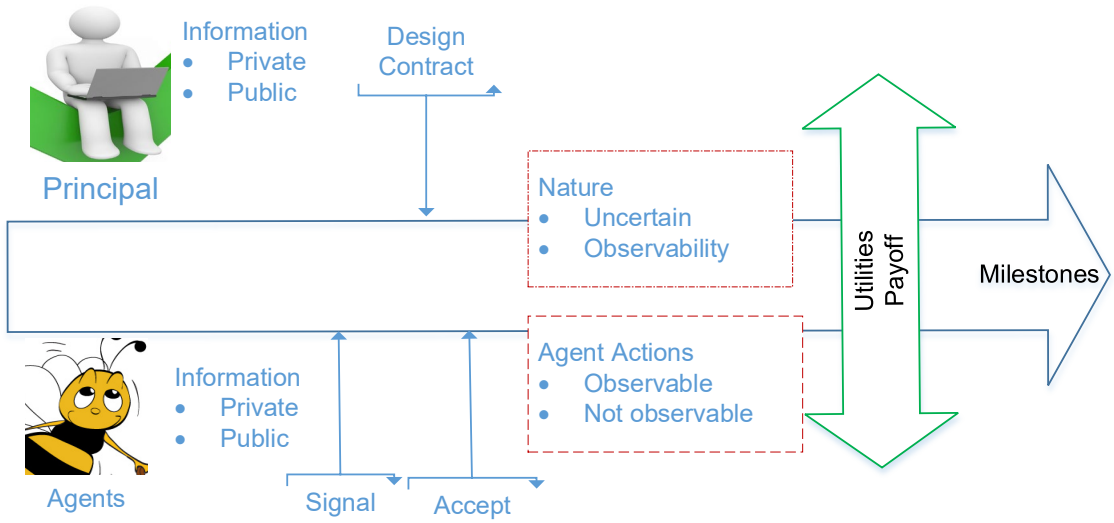


Figure 3.2. Graphical representation of important milestones in the principal-agent model.

This model is an abstraction or simplification. However, it can provide many great insights in forming strategies to find a good agent, a supplier, a client, or a service provider.

2.3 Agent actions

The agent can put in a lot of effort, or little effort, in performing the tasks. A retailer can invest in merchandizing, in training, and in staffing. A transportation provider can invest in their trucks and drivers to provide better service. The different effort can lead to different-level outcomes. It is difficult for the partner to see the effort, though.

In Barilla's case, a distributor can put in a lot of effort and invest to provide best services on Barilla's products to the supermarkets: fast and convenient order processing, giving priority to Barilla's products, on-time delivery, monitoring the market and inventories to avoid stockouts, no errors, and so forth. In doing so, it may help supermarkets to sell more Barilla's pastas and attract more supermarkets. The distributor can also put in less effort on selling Barilla's products, among products from many other suppliers. The contract with each supplier is different and is hidden from Barilla but has strong influence on the supply of Barilla's products.

2.4 Nature or non-controllable factors

Even with the best contract and efforts, the outcomes are impacted by booming or recessing economies, weather, events, disasters, or a pandemic like Covid-19. With the stay-at-home orders during the Covid-19 pandemic, most retail and hospitality businesses were basically stopped. No matter how much effort someone put in, the outcome may not be good. On the other hand, the

game industry, delivery services, video conferencing providers, alcoholic drinks, yeast, bread machines, and campers enjoy a healthy growth.

We call these non-controllable and uncertain factors Nature (N). For a teaching assistant (TA), the outcome of the students served may depend on the instructor, the students in a specific class, the TA room environment, the textbook, and many other factors outside the TA's control.

Some factors in N can be observed by both the principal and agent after their occurrence, such as market fluctuation, natural disaster, and the stockouts at the supermarket. Other factors in N can only be observed by the agent, such as actual students served or the local events or distributors' sick employees that impacted the service.

2.5 Outcomes, pay-offs and objectives of the contract

The total outcome at the end of the contract depends on the contract designed by the principal, the effort by the agent, and Nature. The principal and agent will share the total outcome based on the payoff plan in the contract.

We assume that the principal would design a contract for maximum utility, such as profit, for itself. To achieve this, the principal will design the contract to discourage negative adverse selection and discourage moral hazard.

3. QUANTITATIVE PRINCIPAL AGENT MODEL

The abstractions and assumptions are embedded in the last section. We define the following parameters for the quantitative model:

Π: The expected utility for the principal

π: The expected utility for the agent

e: The cost of the effort by the agent

w: The expected pay off to the agent

O: The expected outcome

$P(e)$: A random variable linking the agent effort, the nature with expected outcome.

3.1 The objective function for the principal

The objective function is to maximize the expected principal's utility. The decision variable is the total expected outcome depending on pay to the agent. For example, if the distributor can sell three truckloads, the reward is a 5% discount on next order. Mathematically, this will be

$$\max_{O(e),P(e)} \Pi = O(e) - w$$

The outcome O depends on the agent effort e and Nature, N. The more effort the agent puts in, the higher the probability for higher levels of outcome. Nature, such as the economic condition or discount by competitors, can also impact the outcome. However, these are uncertain, and therefore will be described by a discrete probability distribution with the mass function

$$Prob[O = O_i | e] = p_{Ri}(e)$$

The subscript *R* designates rational assessment of the probability. We assume the principal had done the market study and will hold a rational assessment of the probabilities based on Nature and effort. For example, at a given effort level *e*, $O = \{5, 4, 3\}$ with probability $p_R = \{0.3, 0.4, 0.3\}$.

3.2 The constraints for the principal

The principal should design the contract so that the expected utility is increasing function and concave function with respect to effort *e*, or if the probability is continuous, we have

$$\frac{\partial \Pi}{\partial e} > 0 \text{ and } \frac{\partial^2 \Pi}{\partial e^2} \leq 0$$

The concaveness is due to two factors. The first is the diminishing rate of return on effort. The second is to increase payoff to the agent to incentivize the agent to put in more effort (see details in next section.)

3.3 The objective function for the agent

The agent would like to maximize its own utility, or

$$\max_{e,P(e)} \pi = w - e$$

The payoff that the principal awards the agent also depends on the agent's effort and Nature, or

$$Prob[w = w_i | e] = p_{Ai}(e)$$

The subscript *A* designates the probabilities as perceived by the agent. The agent may be more risk averse or more risk tolerant then the principal. There are two parts of the agent's utility.

The agent's utility increases with the expected outcomes, which increase with effort. In order to incentivize the agent to put in high effort, the principal should make the reward convex, or,

$$\frac{\partial \pi}{\partial e} > 0 \text{ and } \frac{\partial^2 \pi}{\partial e^2} \geq 0$$

3.4 Constraint for the agents

For the agent to participate, the expected utility must be at least as high as the agent's other opportunities. Or $\pi \geq$ utilities from other opportunities based on whatever effort level they plan to invest.

For the principal to find the best agents in the selection pool, he/she would design a contract to find the highest-effort agents, because the principal's expected utility is a function of the effort and the contract is designed so that

$$\frac{\partial \Pi}{\partial e} > 0$$

There are many possible incentive contracts that satisfy these objectives and constraints, such as discrete, continuous or even non-financial.

4. TWO-LEVEL PRINCIPAL-AGENT MODEL

We will develop a simple two-level quantitative P-A model to find insights in the principal-agent models.

The principal would like to design the contract to maximize its own utility. To do so, it would like to deter the agent who plans to put in low effort to reduce the effect of adverse selection, and to encourage the agent to put in high effort to avoid moral hazard even when the actions are not observable.

Assumptions

1. There are two outcome levels, $O = H = 2$ or $O = L = 1$ containers.
2. There are two agent effort levels, h and l. For example, h = 3,000 for \$3,000 investment for advertisement and training, or $l = 0$ for no special investment.
3. There are two levels of other opportunities, $Other^h$ and $Other^l$. For the same agent, we may have $Other^h = Other^l$.
4. The principal will select two pay-off levels to the agent: w^h when and w^l, depending on the actual and observable outcome H or L.
    - Special case: w^l is base pay, and $w^h = w^l + B$ where B is a fixed bonus.
5. The total revenue is observable by both the principal and agent.
6. The effort level is planned and executed by the agent but not observable by the principal.
7. Nature is observable by both the principal and agent.
8. Both the principal and agent are rational or risk neutral (to be relaxed next).

Principal		
$e \setminus O$	H	L
h	p_{hH}	p_{hL}
l	p_{lH}	p_{lL}

Agent		
$e \setminus w$	w^H	w^L
h	p_{hw^H}	p_{hw^L}
l	p_{lw^H}	p_{lw^L}

Nature will be represented by the probabilities. The high effort to achieve the outcome of *H* and *L* are p_{hH} and p_{hL}, respectively, in the table below on the left. The higher effort will increase the probability of p_{hH}. The agents will focus on its pay, *w*. The probabilities are represented in the table on the right. Based on assumption 8, both principal and agent are rational and risk neutral. Their views on the probabilities are public information and are the same.

Note that the effort the agent put into the tasks is not visible to the principal, based on assumption 6. This is common. Please consider Barilla's case.

4.1 Objective functions and constraint for the principal

The objective for the principal is

$$\max_{e,P(e)} \Pi = \begin{cases} O(h) - w^H = p_{hH}(H - w^H) + p_{hL}(L - w^L), \text{ if } e = h \\ O(l) - w^L = p_{lH}(H - w^H) + p_{lL}(L - w^L), \quad \text{if } e = l \end{cases}$$

The total outcome *O* depends only on the effort level, either *h* of *l*, or Nature, represented by the probabilities. Since the probabilities are given, there are only two possible total outcomes, *O*(*h*), *O*(*l*). The decision variable for the principal is the two-level revenue-dependent pay (wage) to the agent, w^H and w^L, or the share for the agent. The principal would select the agent's pay to ensure maximum utility for itself.

The principal would like to ensure that the expected payoff for the high effort is higher than the expected payoff for the low effort, or

$$p_{hw^H} w^H + p_{hw^L} w^L - h > p_{lw^H} w^H + p_{lw^L} w^L - l$$

The principal will design the contracts based on these objective functions and constraints.

4.2 The objectives and constraints for the agent

The objective function of the agent is

$$\max_e \pi = \begin{cases} p_{hw^H} w^H + p_{hw^L} w^L - h, \text{ if } e = h \\ p_{lw^H} w^H + p_{lw^L} w^L - l, \text{ if } e = l \end{cases}$$

The agent will decide between two levels to suit its need and payoffs. If this represents different agents, the other opportunity may be different. The difference can be incorporated in the constraints.

For high effort,

$$\pi_h = p_{hw^H} w^H + p_{hw^L} w^L - h > Other^h$$

For low effort

$$\pi_l = p_{lw^H} w^H + p_{lw^L} w^L - l > Other^l$$

4.3 An example of a risk-neutral model

When the agent views Nature the same as the principal, the agent is considered risk-neutral. In this case, both probability tables are identical.

Example

A distributor would like to find a new retailer to sell its tools. It will invest \$12 to develop the new dealership if it can find one. The units are \$1,000/week. There are several potential dealers who are interested in the opportunity. If they do not work for this distributor, then they have another opportunity to make *Other* = \$10. The low-level effort costs \$0 (no additional investment) and high-level effort costs \$2 advertisement and training. The revenues are two levels, H = \$60 or L = \$30. Both the distributor and the potential retailers know the market well and assign the same probabilities as follows.

$e \setminus O, w$	$60/w^H$	$30/w^L$
2	0.8	0.2
0	0.4	0.6

Questions

Please find

1. The objective function and constraints for the principal
2. The objective function and constraint for the agent with high effort and low effort
3. The total payoff to be shared by the principal and the agent
4. Draw these constraints on a chart and explain trends of agent utilities
5. What is the desirable region(s) for the principal?
6. The expected utilities for the principal and agent in different regions and trends
7. Insights from this example

Assumption

Already stated.

1. The objective function and constraint for the principal

The objective function for the principal is

$$\max_e \Pi = \begin{cases} 0.8(60 - w^H) + 0.2(30 - w^L) - 12, & \text{if } e = 2 \\ 0.4(60 - w^H) + 0.6(30 - w^L) - 12, & \text{if } e = 0 \end{cases}$$

The constraint for the principal is to attract a high-effort agent and incentivize the agent to put in high effort when performing the task. He would set it so that the utilities for the high effort were higher than the low effort, or

$$0.8w^H + 0.2w^L - 2 \geq 0.4w^H + 0.6w^L - 0$$

Or

$$w^H - w^L - 5 \geq 0 \quad (1)$$

2. The objective function and constraints for the agent

The objective function for the agent

$$\max_e \pi = \begin{cases} 0.8w^H + 0.2w^L - 2, & \text{if } e = 2 \\ 0.4w^H + 0.6w^L - 0, & \text{if } e = 0 \end{cases}$$

Constraint to justify the high effort, the utility for the agent must be at least as much as the other opportunity, or

$$\pi^H = 0.8w^H + 0.2w^L - 2 \geq 10$$

This can be simplified to the nominal form show below. At the equals sign, the agent utility is the same as the other opportunity.

$$w^H + 0.25w^L - 15 \geq 0 \quad (2)$$

The constraint to justify the low-effort participation can be derived similarly:

$\pi^L = 0.4w^H + 0.6w^L \geq 10$, or

$$w^H + 1.5w^L - 25 \geq 0 \quad (3)$$

3. The expected total payoff *O* is a function of the agent's effort only. At the high and low efforts, the expected payoffs are:

$$O(h = 2) = 0.8 * 60 + 0.2 * 30 = 54$$

$$O(h = 0) = 0.4 * 60 + 0.6 * 30 = 42$$

The total pie to be shared by the principal and agent at high effort is 33% higher than the low effort. Therefore, it is important for the principal to design the contract to find and promote high effort.

4. Draw the three constraints on a chart and explain the trends of agent utilities.

The three constraints are illustrated in Figure 3.3.

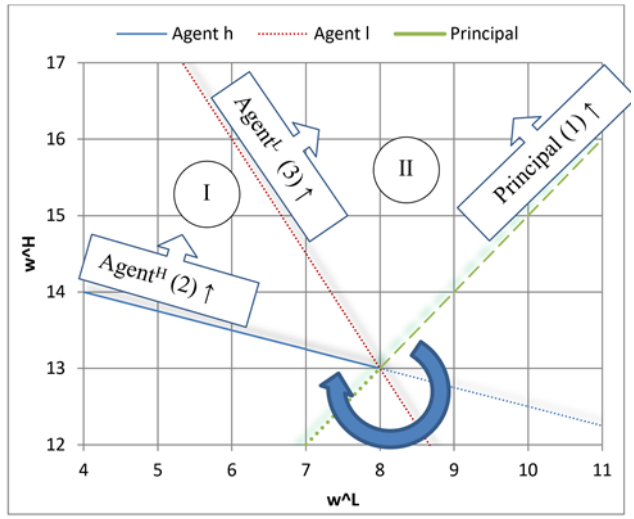


Figure 3.3 The constraint lines and trends for the example.

The principal constraint shown in the dashed line is from constraint (1). On this line, the high effort and low effort would yield the same utilities. If the pay increases in the direction of the arrow, the difference for the utilities between the high effort and low effort increases the fastest. Specifically, the utility for the high-effort agent will increase while the utility for the low-effort agent will decrease.

The high-effort constraint shown in the solid line is from constraint (2). It shows the minimum combination of pay (w^L, w^H) to the high-effort agent. If the pay increases \$1 in the direction of the arrow, its expected pay will increase by \$1.

The low-effort constraint shown in the dotted line is from constraint (3). The explanation is the same as for the high-effort agent.

The intersection among the three lines shows when the utilities for both agents are identical, called for by all three constraints at the equality.

The space swept by the arrow is outside of our interest because at least one of the partners will not participate.

5. What is the desirable region(s) for the principal?

In the region labelled "I," an agent must put in high effort to draw utilities. A low effort will yield lower utilities than the other opportunities.

In the region labelled "II," both high and low effort will yield utilities higher than the other opportunities. The high effort increases the chance of higher outcome and yields more utilities for both the agent and the principal. We will see that a low effort hurts the principal much more than the principal in a good contract.

6. The expected utilities for the principal and agent in different regions and trends

Let us find utilities in some representative locations (w^L, w^H) at (8, 13), (9, 14), (8, 15), (7, 14), and (6, 14).

The offer (8, 13) is on the intersections of three constraints. It is a boundary offer. We can use it as a base point.

$$\max_e \Pi = \begin{cases} 54 - [0.8(13) + 0.2(8)] - 12 = 30, & \text{if } e = 2 \\ 42 - [0.4(13) + 0.6(8)] - 12 = 20, & \text{if } e = 0 \end{cases}$$

It is trivial for the agent because it is the intersection point, which is equal to the other opportunity. You can verify by

$$\max_e \pi = \begin{cases} 0.8 * 13 + 0.2 * 8 - 2 = 10, & \text{if } e = 2 \\ 0.4 * 13 + 0.6 * 8 - 0 = 10, & \text{if } e = 0 \end{cases}$$

This offer of (8, 13) only makes a difference for the principal. It does not matter for the agent. Therefore, there is no point for the agent to put in high effort. Moral hazard will occur, and lead to low utility of 20 for the principal.

The offer (9, 14) will lead to utilities

$$\max_e \Pi = \begin{cases} 54 - [0.8(14) + 0.2(9)] - 12 = 29, & \text{if } e = 2 \\ 42 - [0.4(14) + 0.6(9)] - 12 = 19, & \text{if } e = 0 \end{cases}$$

$$\max_e \pi = \begin{cases} 0.8 * 14 + 0.2 * 9 - 2 = 11, & \text{if } e = 2 \\ 0.4 * 14 + 0.6 * 9 - 0 = 11, & \text{if } e = 0 \end{cases}$$

Since this is on the principal's incentive curve, it means there is no incentive for the agent to put in high effort, since low effort will lead to the same pay. Since both low pay and high pay are 1 unit above, the expected utility for the agent will only be 19, even worse than (8, 13).

At (8, 15)

$$\max_e \Pi = \begin{cases} 54 - [0.8(15) + 0.2(8)] - 12 = 28.4, & \text{if } e = 2 \\ 42 - [0.4(15) + 0.6(8)] - 12 = 19.2, & \text{if } e = 0 \end{cases}$$

$$\max_e \pi = \begin{cases} 0.8 * 15 + 0.2 * 8 - 2 = 11.6, & \text{if } e = 2 \\ 0.4 * 15 + 0.6 * 8 - 0 = 10.8, & \text{if } e = 0 \end{cases}$$

At this point, the utilities for both the high effort and low effort would see an increase in agent's utility, although the increase in utility for high effort is twice that for low effort. However, it is

difficult for the principal to tell if the agent plans to put in high effort or low effort. This leaves room for adverse selection. In addition, if the agent plans to put in high effort, it may actually put in low effort when other things come up, as in the real world. Then, moral hazard will occur, since the utility is still better than the other opportunity.

At (7, 14)

$$\max_e \Pi = \begin{cases} 54 - [0.8(14) + 0.2(7)] - 12 = 29.4, & \text{if } e = 2 \\ 42 - [0.4(14) + 0.6(7)] - 12 = 20.2, & \text{if } e = 0 \end{cases}$$

$$\max_e \pi = \begin{cases} 0.8 * 14 + 0.2 * 9 - 2 = 10.6, & \text{if } e = 2 \\ 0.4 * 14 + 0.6 * 9 - 0 = 9.8, & \text{if } e = 0 \end{cases}$$

The utility for the low effort is lower than the other offer. Therefore, the agent plans to put in low effort and will not sign the contract. This will prevent adverse selection. The utility for the agent plans to put in high effort is higher than the other option, and therefore, can sign the contract. This is what the principal wants. After signing the contract, the agent has incentive to sustain the high effort to ensure the utility is higher than the other offer. This will prevent moral hazard! Therefore, the offer of (7, 14) is excellent. The expected utility for the principal will be 29.4, only 0.6 units before the maximum possible.

(7, 14) is the best integer offer in this problem. As discussed earlier, all solutions in Region I are good. The other integer offers shown in the figure are (6, 14), (6, 15), (5, 14), (5, 15), (5, 16), (5, 17), (4, 15), (4, 17)…. However, those offers lead to lower utilities for the principal. In addition, the real problems may not be linear. At some point, the agent may not want to take the risk.

7. Insights from this example

You can gain some insights from this example.

1. If (w^L, w^H) are in the region I, an agent must plan to put in high effort, and actually put in high effort to achieve utilities higher than the other opportunities. This will deter the adverse selection and moral hazard, which will increase the probability of higher outcome. This will align the incentives between the principal and the agent.
2. Within region I, the further away from constraint (1), the bigger the gap between the low pay and high pay, and the stronger the incentive for the agent to plan and to apply high effort. The further away from constraint (2), the higher expected utilities for the agent. The further away from constraint (2), the bigger the gap between the utilities of high effort and low effort, and therefore strong deterrent for adverse selection and moral hazard. The stronger incentive for the agent is due to lower low pay. It will also increase the principal's utilities due to lower low pay. In this simple model, the extreme solution would be (0, 15).
3. If the offer is in region II, both high and low effort will yield higher utilities for the agent than other opportunities. It will not deter adverse selection. An agent plan to put high effort may not maintain its high effort because a lower effort will still yield utilities higher than the other opportunities.

4.4 Risk-averse agent

A risk-averse agent is less optimistic on the revenue prospects. It can be modeled with lower probabilities to the high revenue and higher probability of low revenue. From the above example, in the same market with the same probabilities viewed by the principal, a risk-averse agent may consider the probabilities as

$e \setminus R(e)$	50	30
2	0.8−0.1 = 0.7	0.2+0.1 = 0.3
0	0.4−0.2 = 0.2	0.6+0.2 = 0.8

In this case, the principal's objective function and constraints stay the same. However, the agent's objective function and constraints will change. The objective function for the agent will be

$$\max_e \pi = \begin{cases} 0.7w^H + 0.3w^L - 2, \text{ if } e = 2 \\ 0.2w^H + 0.8w^L - 0, \text{ if } e = 0 \end{cases}$$

The principal's incentive constraint depends on if the principal knows the agent's attitude toward the risk. If the principal does not know the agent is risk averse, the constraint stays the same as in (1). The high effort constraint is

$0.7w^H + 0.3w^L - 2 \geq 10$, or

$w^H + \frac{3}{7}w^L - \frac{120}{7} \geq 0$

The low effort constraint is

$0.2w^H + 0.8w^L - 10 \geq 0$, or

$w^H + 4w^L - 50 \geq 0$

These constraints are shown in Figure 3.4.

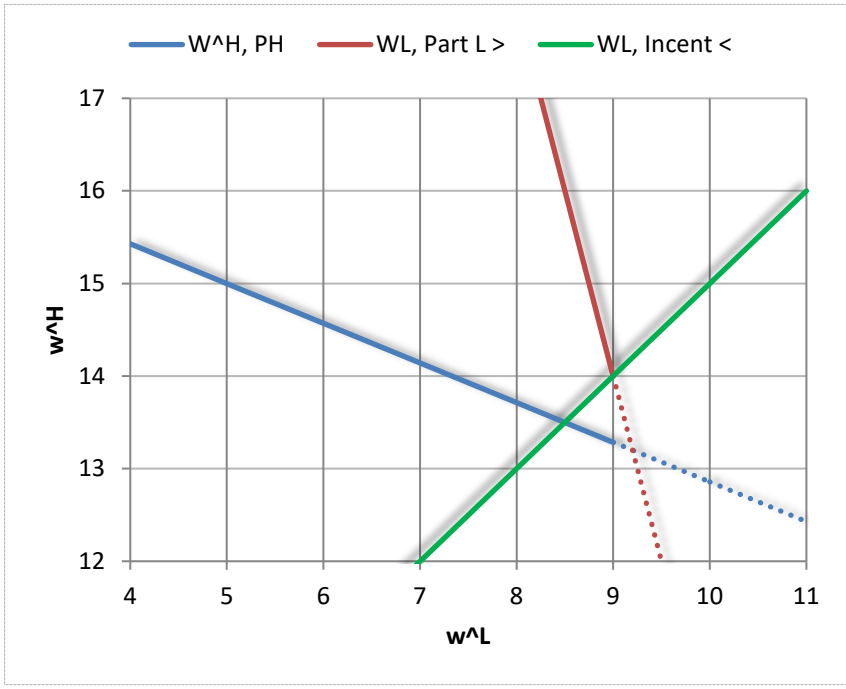


Figure 3.4. The constraints with risk-averse agent.

Note that the agent's constraints intersection point moves to the right and slightly up. The intersection point of the agent's constraints is no longer on the principal's incentive constraint. Although the agent's actual utility stays the same, the principal has to offer higher pay for the principal to accept. Most of the increase must be the low pay. The reader can also consider what if the principal has detected that the agent is risk averse, how to set principal's incentive constraint.

4.5 Agents with different levels of other opportunities

Different potential agents may have different levels pay opportunities. The same agent may have different pay opportunities if planned for a different level of effort. If that is the case, the right-hand side of the participation constraint will change. The constraints may have multiple intersection points. You can apply the same logic to find the best regions for the principal in the contract design.

4.6 Effort observability

So far, we assume that the principal cannot observe the effort that the agent put in, stated in assumption 6. However, it is possible for the principal to observe some efforts the agent put in directly or indirectly. Many firms hire secret shoppers to observe the services provided by the retailers. Many suppliers send their personnel to work in the third-party logistic providers. They can observe the effort the service providers place on their products. Honda set high-quality

standards to their suppliers with very low part per million counts, to the level that many suppliers cannot achieve even if they want to. Honda would send their engineers in to help. These are direct observations.

The suppliers can also observe the agent's effort indirectly through something visible during the contract interval. For example, Barilla can negotiate with the distributors and supermarkets to find daily sales of their products, from which estimates of the effort the agent put in can be made.

4.7 Two-level principal-agent model summary

Principal-agent models provide excellent insights into various partnership situations. In a supply chain, the asymmetry in information, objectives, and incentives can lead to adverse selection and moral hazard that can hurt partners. It is important to understand the asymmetry and design the contract to use the incentives to create synergy between the partners and to achieve win-win situations. A simple two-level principal-agent model was used to illustrate the dynamics due to the information asymmetry. It is important for the engineers and managers to understand goals of the people and organization, identify the sources of adverse selection and moral hazard, and design the solutions to align the incentives to achieve better efficiency among partners.

In practice, incentives can have many levels. The incentives can be in many dimensions, such as sales-based discounts, technical help, return policies, and priority treatments. Incentives can also be in the social dimension, such as recognitions or awards. Modeling for Nature can also be continuous probability functions. For one principal, there can be many agents. There are different relationships among the agents. There can also be multiple principals with mutual agents. These will complicate the models. However, the signaling, adverse selection, moral hazard, information asymmetry are still the main factors to be considered in designing incentive structures.

5. TOWARD VENDOR-MANAGED INVENTORY

With the information presented in Chapter 2, Collaboration, the Barilla case and the HBR article on aligning supply chain, we can now discuss more specifics on vendor-managed inventory (VMI).

The JIT distribution proposed by the logistics manager in Barilla is an effort toward VMI. In supply chain, inventory is always a major concern. Most inventories are stocked for order, or speculative (see Cisco's story in the HBR article). They take funds, occupy space, cost money to transport, and are subject to many risks. Collaboration between supply-chain partners has the potential to greatly reduce the total inventory risk levels with some form of VMI. The basic idea of VMI is to manage the inventory across the partners to provide best availability with best response and minimum stockouts.

VMI will always need a contract to bind the responsibilities of the participating partners. The principal can be any of the players along the chain who will design the contract. The others are agents. The principal can follow the solution steps, and design the contracts to align the objectives, to reduce adverse selection and moral hazard.

5.1 The key characteristics of VMI

There are many forms of VMI. One extreme is case is Coca Cola, which sells their products in certain supermarkets. The supermarkets allocate some space for Coca Cola projects. Coca Cola

manages the display, the assortment and the replenishment, and owns the inventory. When a customer checks out products at the cash register, the transaction is stored and later passed on to Coca Cola at the end of the day. The supermarket will send a percentage of the sales to Coca Cola based on the contract. The supermarket does not own the inventory, stockout, or any risk directly.

The JIT distribution proposed by the logistics manager is an effort for the vendor Barilla to manage the inventory at the distributor, and can be considered an effort toward VMI. The distributor still stores the inventory and owns the inventory. It is up to the producer and the distributor to figure out the liabilities of stockouts, price drop, or other "risks" involved.

Another example can be that the store passes the sales information to the supplier for the supplier to manage the time and quantity of replenishment. The store owns the inventory after the replenishment is completed based on the terms in the contract. If the goods become obsolete, the burden is on the store. The storage can also be at a third party's distribution center, such as a third-party logistics (3PL) provider. The ownership depends on the contract.

There are a number of key factors in VMI systems.

1. Visibility: What information is visible electronically between the partners? Modern retailers have access to real-time demand and inventory information. All channel partners can benefit from some level of visibility in demand and inventory. Suppliers can use the information to plan their production and inventory. The sharing of complex information may need certain technology such as EDI that allows different applications by different partners to put the data, such as SKU, quantity, and date, into proper places in the partner's respective applications automatically.
2. Ownership: Who owns the inventory from what time? Ownership comes with the risk: overstock, understock, damage, shrinkage, price drop, obsolescence, and so forth.
3. Location: Where are the inventories stored? The more downstream the storage location, the faster the response to the end-consumer (economies of speed). However, the more downstream, the more committed the inventory is to a specific market. The more downstream, the smaller the quantity, which means fewer economies of scale, from pallets to cases to pieces. The retail stores are normally located at expensive commercial space, and have good displays with good lighting and expensive HVAC systems. Distribution centers can be in a low-cost area without fancy lighting or air conditioning. The inventory can also be stored at a third-party logistics provider's location.
4. Trigger: What will trigger replenishment? A simple trigger is time, for example, every week. A more sophisticated trigger can be the inventory position, such as Kanban or empty container. It can also be real-time electronic monitoring using sensors. If the replenishment time and quantity are by inventory position, it is a pull system. If the replenishment is based on time with replenishment quantity based on forecast, such as in the case of JIT distribution the logistics manager in Barilla proposed, it is a push system. There can be hybrids in which the trigger is time but the quantity depends on the inventory position.
5. Security: How is security handled? Pilferage can be problem for many products, such as razors, toys, and certain electronics. In distribution centers, certain products can be stored

in cages. In stores, certain product displays can be made difficult to steal from (such as taking a long time and effort to take one). The security measures will increase cost.

6. Status change: Have things altered? When the partners signed a contract, both sides may be motivated. However, the situation can change. For example, the agent may sign a more lucrative contract and shift their priority to others.
7. Trust: Can the principals trust their personnel? Without good observability and trust, the principals can request their own personnel at the agent's sites. The principal must find ways to achieve this such as Honda send their engineers to their supplier's site to help with quality assurance.

5.2 Some examples of VMI

Here are three examples of what some may call VMI. Please determine their types based on visibility, ownership, location, and trigger, and the benefits of collaboration, shown in Table 3.1.

Table 3.1. Possible practices some call VMI.

Examples	Contract Arrangements	Characteristics
1	1. The distributor shares daily sales for each SKU with the producer. 2. The distributor orders from the supplier based on its inventory positions, prices, and markets. 3. The producer produces to stock and deliver based on the orders.	1. Visibility 2. Ownership 3. Location 4. Trigger 5. Benefits to each partner
2	1. The distributor share daily sales and inventory position for each SKU. 2. The producer plans the production and delivery based on inventory position on a fixed schedule.	1. Visibility 2. Ownership 3. Location 4. Trigger 5. Benefits to each partner
3	1. The distributor share hourly sales and inventory positions for each SKU. 2. The producer plans the production and delivery based on inventory position.	1. Visibility 2. Ownership 3. Location 4. Trigger 5. Benefits to each partner

5.3 Principal agent model and VMI

For each scheme, there are costs and benefits for different partners. The principal should determine objectives, potential adverse selection, moral hazard, and design the contract with incentives to benefit all partners. The incentives can be in pricing, refund, and improved service. These incentives can also be in levels, are not restricted to two levels.

6. PRINCIPAL AGENT MODEL AND HUMAN NEEDS

The objectives of for-profit firms include profit maximization. Even nonprofit organizations such as hospitals strive for operational excess, so that they can invest to improve their services. When the objective of profit or operational excess is aligned with human needs, the focus on financial utility is sufficient. Supply-chain professionals should consider strategies to align the objectives.

6.1 Strategies of the conference center waste management project

In the waste management in the conference center example in Chapter 2, the senior design team linked the waste haul charges to diversion rate to nudge the conference hosts to diverge more from the land fill. The waste haul charges are observable by the conference center. The two-level principal-agent model can provide qualitative and quantitative insights into the adverse selection and moral hazard in the design of the charging system. The conference hosts would be conscientious when hiring and dealing with their contractors to reduce the waste-haul charges. They can even pass on the extra charge to the contractors to give contractors a direct incentive to increase the diversion rate.

The senior design team also incentivizes the conference host using awards associated with social and environmental utilities based on observable outcomes.

6.2 Singapore's strategy to align finance and human's health needs

In Singapore, the tax-funded public programs only pay for about one-quarter of the total healthcare, based on the Pacific Prime [1]. For almost all services, the patients must pay a portion from a health savings account called Medisave, a mandatory savings account deducted from payrolls. There is also a public health insurance, Medishield, that pays a portion of the more expensive charges. The government also provides subsidies for certain expensive but necessary treatments in the public hospitals. The self-pay helps the patient to align the financial utilities and health utilities, instead of "care" utility.

The universal coverage Medishield and subsidy avoids the adverse selection from the patient, and the significant pay from Medisave by the patient also encourages people to maintain health and avoid the moral hazard of over-consumption of services or medicine.

The Singaporean government requires hospitals to publish their prices and doctors to reveal their benchmark fees. The pricing transparency reduced the information asymmetry. The payments from all parties are predetermined. Therefore, there is no waste in the administrative costs after the treatment, which is protracted in the United States.

From WHO health data, we can subtract the healthy life expectancy from birth, from the life expectancy from birth, to get the expected number of unhealthy years, shown in Figure 3.5. By this measure, the United States has the best performance. Given the incentives in "care," this is not surprising. In Singapore, there are strong incentives in "health," which also show with the

fewest expected unhealthy years. Comparing Figures 1.2 and 3.5, we get great contrast of the population health and healthcare systems between the United States and Singapore. The principal-agent model provides great insight for this outcome.

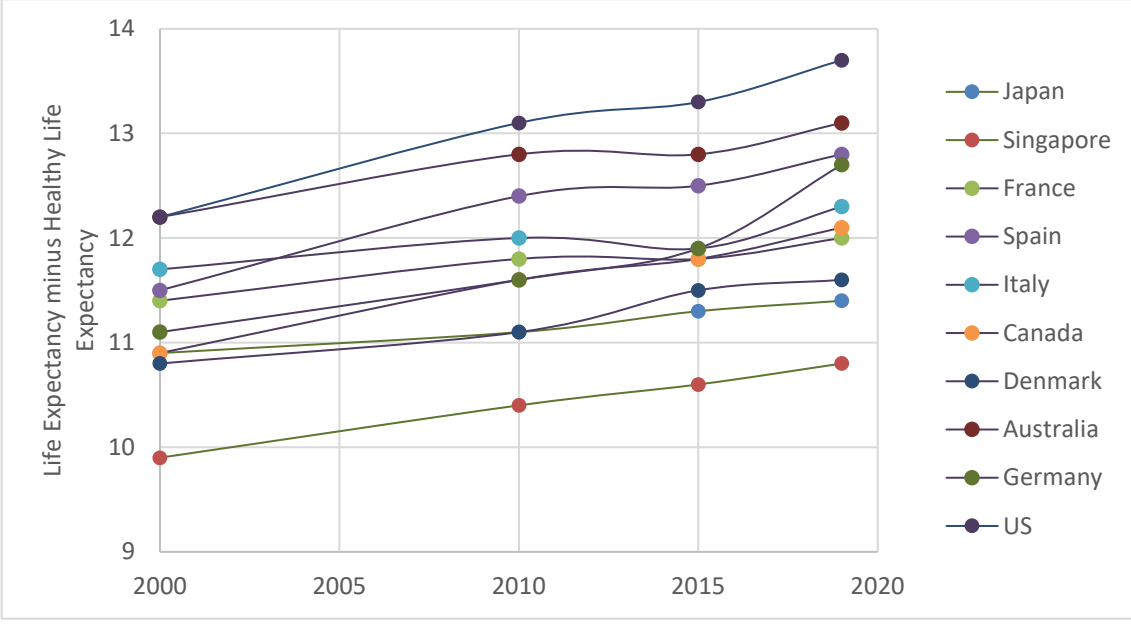


Figure 3.5. The expected number of unhealthy years in select countries, from 2000 to 2019.

6.3 A U.S. example to align finance with human health needs

First of all, the human need is for health. The "care" becomes important only if humans lose health. However, the financial incentives in healthcare are in the "care," which may not always be aligned with health. Kaiser Permanente in many states in the United States is an integrated health service provider that integrate finance with health services. It collect premiums, and work with its own health services. Since Kaiser has to pay for the service charges using the premiums collected from the patient, it has an incentive for health because healthy people can avoid costly healthcare expenses. A major characteristics is that Kaiser spends much more effort to promoate healthy living. Another characteristics is its simplicity among different enties and the patient. However, Kaiser is not perfect, as in any other systems, and has certain shortcomings such as potential long wait times and lack of choices. But health performance has been good in many dimensions.

6.4 Issues in the U.S. healthcare system

In the United States, there are many players in healthcare system, shown in Figure 1.2 in Chapter 1. First of all, the service providers and insurance companies provide services for treatment, which may not be aligned with the utility of health. From that point of view, the U.S. healthcare system does very well, as shown in Figure 3.5: the system kept people alive even when they lost health. In terms of cost, you can learn about the interesting drug supply chain and pricing from a six-minute video, "How Drug Prices Work" [2]. It shows how the pharmaceutical firms, distributors, pharmacy benefit managers (PBMs), and insurance companies collaborate but lead to higher prices for consumers.

Each organization in the complex network of a healthcare system has its own financial and performance objectives. Their objectives may not always be aligned with patients objectives in health. The objective misalignments, the information asymmetry, lack of transparency, vendor-induced over-testing and over-treatment, patient adverse selection, and moral hazard of all parties, leads to the high cost and low performance in the system.

One of the senior design teams was working with a hospital on a procedure to treat severe cases of epilepsy: to separate the two sides of the brain. The procedure generates more relative operational excess than other procedures. However, the hospital only performs very few such procedures each year, mostly due to the lengthy and slow approval process controlled by the regulations. From a financial perspective, the objective can be to speed up the approval process. From a healthcare performance perspective, separating two sides of the brain has major health implications. The regulation in place is to ensure the health performance of the procedures.

4. COMPETITION IN SUPPLY CHAIN

Supply meets demand at the market. In most markets, there are multiple players. Each player must compete to maximize their utilities, such as profit. We will now look at the other three competitive forces in Michael Porter's five competition forces. We will then focus on the competition in a specific market. In a specific market, the competitors compete with similar products by gaining market share or setting prices. The competitors can also compete by product or service differentiation. The competition with market share can be modeled in Cournot competition, in which the supplier plans the supply quantity; the supply quantity will determine the price; and the quantity, price, and cost together will determine the profit. The competition with price can be modeled as a Bertrand competition, in which the supplier first determines the selling price. The price will determine the supply quantity and the profit in the market. In both supply quantity and price competition, cost-cutting and efficiency improvement can increase the profit margin of similar products or services.

The competitors can also compete through differentiation. Differentiated products can command higher prices and increase profit. A Bertrand competition with product differentiation can capture this market dynamic.

We finally discuss the strategies a firm can use based on all of the factors in different types of markets.

1. PORTER'S OTHER THREE FORCES IN COMPETITION

Figure 2.1, in Chapter 2 Collaboration in Supply Chain, shows that there are other market forces. We will discuss them here.

1.1 Threat of new entrants

In any market, there is always a possibility of new entrants. Calculators have become commodities. Many firms are able to supply them, even in emerging economies. The supply cost of simple products from emerging economies is often lower than in the advanced economies. The barriers to entry to supply different products vary. For example, the investment cost of a well-designed game can be low. The investment to build a facility to assemble production cars or semiconductors can be very high. The higher the initial investment, the higher the technical contents and the higher the barrier to entry. The incumbents may also enjoy the supply side of the economies of scale, the customer base on the demand side, cost of switching to new suppliers, access to distribution channels, restrictive government policies, and other factors such as market growth rate. The existing competitors may also form alliances to suffocate the new entrants, which could cross the legal line.

1.2 Threat of substitutes

Products can be substituted. Every smartphone now comes with calculator functions sufficient for many people. Block Busters was replaced by Netflix, which itself is being challenged by various streaming services. Combustion-engine-powered cars are being replaced by electric-powered cars. The point-and-shoot camera was replaced by smart phones, for the most part.

Certain brick and mortar stores are being replaced by e-commerce. Taxis and certain delivery services are being replaced by shared rides or shared delivery services.

The threats of the substitute are high if they offer lower cost or higher performance, or the client's switching cost is low, or the new products also offer additional desirable features in addition to the ones offered by the current suppliers.

1.3 Rivalry among existing competitors

This is most relevant to the discussions in this book. The competition is fiercer if:

1. There are more competitors. We can quantify this in the Cournot competition model. The more competitors in the same market, the lower the profit margin, and the smaller market share for each competitor.
2. The total market growth is slow. This is similar to a zero-sum game and leads to cut-throat competition.
3. The exit barrier is high. If a firm exits a market, its existing infrastructure may not be useful for the new market. You can consider a software game provider moving toward a new game versus an oil refinery or automobile assembly.
4. The competitors compete on the same dimensions, such as price, speed of delivery, calorie contents, and sweetness. The worst case with respect to profitability is price competition. All the models in this chapter will show that.
5. The products are nearly identical.
6. The fixed cost is high and the marginal cost is low. We will see this more in the Chapter 7, Economies of Scale.
7. Capacity expansion involves large capital investments in large increments.
8. The product shelf-life is short. There is pressure to reduce prices before the products expire.
9. Differentiation can be good to increase the market itself.

2. MONOPOLY MAXIMUM PROFIT SUPPLY QUANTITY

2.1 Reverse demand model

The market sets the prices between the supply and demand. The higher the supply, the stronger the downward push for price. The supply and demand of a product are complicated, especially in modern society with so many options of where and what to buy for what purpose. However, some simple models can provide great insight into effective competition. Most models for competition are based on the linear decreasing supply-price model, which economists refer to as inverse demand curve.

Definitions

1. q: supply quantity
2. p: price of one product or service, a function of the supply quantity
3. P: maximum possible price
4. Q: maximum possible supply or demand

Assumptions

1. The price decreases linearly with supply quantity.

We adopt one canonical form of a linear function in the form $y = Ax + B$. If we substitute with p, q, B, P we have

$$p = P - Bq$$

The graph is shown in Figure 4.1.

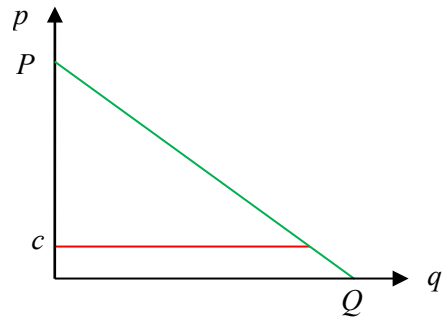


Figure 4.1. The price-supply and cost curves.

When $q = 0, p = P$. When $p = 0, q = Q, B = \frac{P}{Q}$.

Example

A company supplies high-fashion winter boots. The estimated maximum demand for one design is 200,000 when the company would give it away. The cost to supply a pair is $120. The maximum price to have any customers is $200. Please find the supply-price model.

Solutions

Assume that both the linearity and constant costs assumptions hold. In practice, they are not true. More discussions later.

When $q = 0, p = P = 200$. When $p = 0, p = 0, B = \frac{P}{Q} = \frac{200}{200000} = 0.001$. The reverse demand model is then

$$p = -0.001q + 200$$

2.2 Monopoly maximum profit model

When a firm monopolizes a product or service, there is no competition. In the United States, monopoly has been rare. When AT&T reached monopoly, the government stepped in to break it apart. Another example is Standard Oil. There has been talk of the monopolies of Google and

Amazon. There were also talks of a Windows monopoly for a while. However, Apple posts significant competition. There was AT&T's bid to acquire T-Mobile in 2011. T-Mobile at the time was the third-largest provider after Verizon and AT&T. The U.S. government threatened to break up the deal, and the merger never occurred. Office Depot bought Office Max in 2011 and faced scrutiny. Staples offered to acquire Office Depot on February 4, 2015. The FTC has been studying if the merger will lead to monopoly. By March 31, the FTC was still requesting information from Office Depot. In May 2016, Staples called off its attempted merger.

French economist Cournot concluded that firms maximize profit by varying supply quantity. The price will be determined by the market once the supply quantity is fixed; and the supply quantity, cost, and resultant price in the market determine the profit.

Definitions

1. *c:* marginal cost of producing another product, a constant

Assumptions

1. The reverse demand model is accurate
2. The supplier is to maximize the profit
3. The supplier decides on planned supply quantity
4. The marginal cost does not change with supply quantity

Modeling

Using the price-supply model, we can write the profit function

$$\pi(q) = q(p - c)$$

Substituting p by P - Bq in the reverse demand model, we have

$$\pi(q) = q(P - Bq - c)$$

The profit is the rectangle area shaded green in Figure 4.2. When q is low, the area is thin and tall. As the quantity increases, the area becomes shorter but wider, with increasing total area. When q is very high, the area becomes wider but also shallow, with decreasing total area. There must exist an optimum supply quantity.

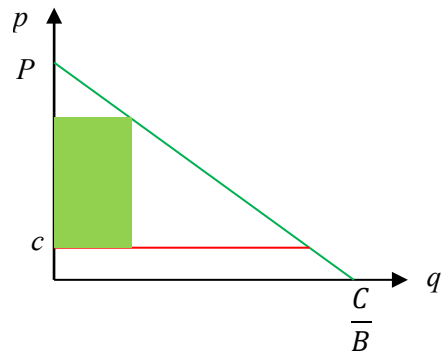


Figure 4.2. Profit based on the price-supply curve.

We can find the profit maximizing supply quantity by the first-order condition.

$$\frac{d\pi(q)}{dq} = -2Bq + P - c \Rightarrow 0$$

We can find

$$q^* = \frac{P-c}{2B}$$

and

$$p(q^*) = \frac{P+c}{2}$$

and

$$\pi(q^*) = q^*(p^* - c) = \frac{(P-c)^2}{4B}$$

These results are shown in Figure 4.3.

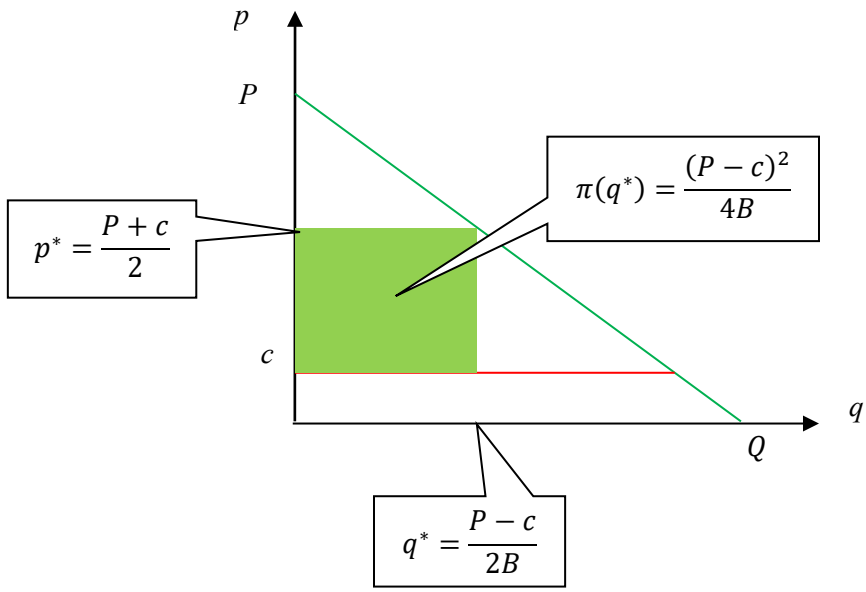


Figure 4.3. Solutions for the monopoly supply quantity, price and profit.

Observations

- The firm monopolizing the market can set the prices through supply quantities.
- The profit-maximizing supply quantity is half of the breakeven supply quantity.
- The profit-maximizing price is the average of maximum possible price and cost (P, c).
- If you can reduce the cost, the profit maximizing supply quantity increases, the price decreases linearly, and the profit increases at the square of the difference between maximum price and cost.

Example

A paper mill makes a unique type of paper. The marginal cost for making this type of paper is c = \$400 per case. The maximum possible price is \$1,000, and the maximum market is 2,000 per month. What is the profit-maximizing supply quantity, the price and profit?

Solutions

We first need to find the parameters from the equation:

When $q = 0$, $p = P = 1000$. When $p = 0$, $B = 0.5$, therefore

$$p = 1000 - 0.5q$$

$$q^* = \frac{P-c}{2B} = \frac{1000-400}{2*0.5} = 600$$

$$p^* = \frac{P+c}{2} = \frac{1000+400}{2} = \$700$$

$$\pi^* = \frac{(P-c)^2}{4B} = \frac{(1000-400)^2}{4*0.5} = \$180{,}000$$

The supply quantity, its resulting price based on the linear model, and profit based on the Cournot model are shown in Figure 4.4.

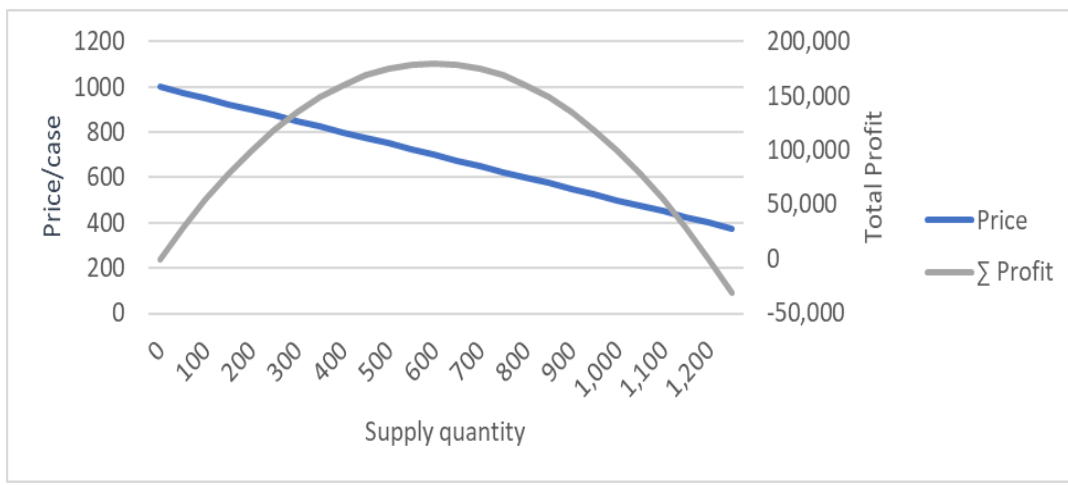


Figure 4.4. Price and total profit as a function of the supply quantity.

Discussions

In addition to what was noted in the observations, you can verify the effectiveness of cost reduction. If you can cut down the cost by 50% from $400 to $200, the supply profit-maximizing supply quantity is increased to 800, the price will go down to $600, but the profit will increase to 3,200,000, or 77.8%!

3. COURNOT COMPETITION IN DUOPOLY

In the real world, multiple suppliers compete to supply products or services. Let us start with duopoly. There are markets where two suppliers dominate: Boeing and Airbus, Coke-Pepsi, UPS-FedEx, Home Depot-Lowes, Canon-Nikon, AT&T-Verizon.

In Duopoly, each one has brand reputation, and a robust supply-chain financial position through years of growth, mergers, and acquisitions. As a result, it is difficult to eliminate any of them. It is also difficult to merge again, because the government may intervene using antitrust laws. Since each supplier supplies a major part of the market, their supply quantity can influence the price significantly.

3.1 Modeling and solutions

Additional assumptions to the monopoly model

- Firms 1 and 2 supply identical products in the same market.
- The supply costs per unit are the same and transparent for both firms.
- The combined supply quantity and reverse demand model determines the price.

Modeling and solution

Based on these additional assumptions, we have the reverse demand model for duopoly

$$p = P - B(q_1 + q_2)$$

Profits or payback functions for each player are

$$\pi_1(q_1, q_2) = q_1(p - c) = q_1[P - B(q_1 + q_2) - c]$$

$$\pi_2(q_1, q_2) = q_2(p - c) = q_2[P - B(q_1 + q_2) - c]$$

Applying first-order condition using partial derivatives, we have

$$\frac{\partial \pi_1}{\partial q_1} = P - B(q_1 + q_2) - c - Bq_1 = P - c - 2Bq_1 - Bq_2 \to 0$$

We can find

$$q_1^* = \frac{P-c}{2B} - \frac{q_2}{2} \qquad (1)$$

Similarly,

$$q_2^* = \frac{P-c}{2B} - \frac{q_1}{2} \qquad (2)$$

Equations (1) and (2) are called best-response functions when the quantity from the opponents is known. In reality, it is not known.

Observations:

- The first term is the maximizing quantity for the monopoly.
- The second term is half of the competitor's quantity. This is interesting in that the profit-maximizing supply quantity for one supplier depends on its competitor!
- If the competitor supplies more, a firm should supply less, because both are bound by the same decreasing price, or a race to the bottom.

3.2 Playing the Cournot game

Let us play a game using the parameters in the special paper supply monopoly example. Recall the first team in (1) and (2). The profit maximizing quantity for monopoly is q^* = 600. Anything above 600 will lead to loss with just one player. Therefore, it is not a good idea. In fact, you want to be significantly less than 600 because the other supplier also supplies.

In class, we can use a TurningPoint application to play this game. Online, we can use MobLab to play this or other similar games. We can limit the choices to multiple of 10s.

The preparations:

1. Form a pair to compete. Remember your competition's name.
2. Find a piece of paper or start a spreadsheet, as shown below.
3. Create a table on paper or computer with the following columns.
4. Record the date, your name, and your competition's name.
5. Review the monopoly optimum quantity and equations (1) and (2)

Rd	q_1	q_2	$q_1 + q_2$	$p = 1000 - 0.5(q_1 + q_2)$	Your π $(p-400)q_1$	Opponent π $(p-400)q_2$
0						
1						
2						
3						

The game is a different type of optimization. Your payoff depends on your competition's decisions!

3.3 Nash equilibrium

This creates a challenge to optimization thinking. You have to try to understand your opponent or make assumptions about your opponent. The Nobel Prize laureate on economics John Nash developed the concept of the Nash equilibrium (NE). The Nash equilibrium for multiple decision-maker problems in games is analogous to the concept of "optimum solution" with single decision-maker problems. With NE, no rational player has an incentive to move away from the decision point after considering the opponent's decisions. You may already get some understanding in the game you played. If not yet, you will see soon.

If we assume

1. Both players are equally rational, and
2. Both players have equal position in the marketplace

Then, they would set their supply quantities the same. This means we can find NE by setting $q_2^N = q_1^N$,or

$$q_1^N = \frac{P-c}{2B} - \frac{q_1^N}{2}$$

We can solve this single-variable linear equation and find

$$q_1^N = q_2^N = \frac{P-c}{3B}$$

$$p = \frac{P+2c}{3}$$

$$\pi_1^N(q_1^N) = \pi_2^N(q_2^*) = \frac{(P-c)^2}{9B}$$

Example

Noticing the special paper the firm monopolized in the last example, or firm 1, was so profitable, another paper producer, or firm 2, developed a similar product and started to market and get ready to compete with the current supplier. The firms set their prices each week. What should be the supply quantity? We do not consider the effect of first mover and second mover, which is studied by Stackelberg.

1. If both suppliers are rational, what should be their supply quantity, resulting price and profit for each?

2. Consider the game, and the monopoly situation, and determine what are the profit-maximizing supply quantity, resulting price, and profit for each?
3. Please observe the solutions in part 1 and 2 and explain why the suppliers do not adopt a profit-maximizing supply quantity? How much do they stand to lose by doing so?
4. What is the benefit of the Nash equilibrium?
5. What are some of the practical considerations?

Assumptions

1. All the assumptions for the price-supply model are satisfied.
2. All the assumptions for the Cournot competition hold.

Solutions

1. Based on the above assumptions, we can find the price-supply function as

$$p = 1000 - 0.5(q_1 + q_2)$$

The profit or payback functions for players 1 and 2 are:

$$\pi_1(q_1, q_2) = q_1(p - c) = q_1[1000 - 400 - 0.5(q_1 + q_2)]$$

$$\pi_2(q_1, q_2) = q_2(p - c) = q_2[600 - 0.5(q_1 + q_2)]$$

Taking the first-order condition, we have:

$$q_1^* = 600 - 0.5q_2$$

$$q_2^* = 600 - 0.5q_1$$

At NE, we have $q_1^N = q_2^N$, or

$$q_1^N = q_2^N = \frac{P-c}{3B} = 400$$

Note together that they supply 2000 more than the monopoly supplier.

$$p^N = \frac{P+2c}{3} = 600$$

Note that this price is $100 lower than the monopoly price. Good for the customers! The profit:

$\pi_1^N(q_1^N) = \pi_2^N(q_2^*) = \frac{(P-c)^2}{9B} = 80{,}000$

The combined profit is \$160,000, \$20,000 less than the profit from the monopoly supplier.

2. Profit-maximizing supply quantity, price, and profit

If you are observant, you may notice that if each firm supplies half of the monopoly quantity at 300 each, or $q_1^* = q_2^* = 300$, $p = 700$, and $\pi_1^* = \pi_2^* = 90{,}000$. The profit for each will be \$10,000 higher than at NE!

3. The reason for why not adopting maximizing not NE

If supplying half of the monopoly quantity yields 12.5% higher profit, why did the firm not adopt that quantity?

Let us start with what will happen if firm 1 supplies a little more, say 350? The total supply now will be 650. The price will be \$675.

$\pi_1 = (675 - 400) * 350 = 96{,}250$

$\pi_2 = (675 - 400) * 300 = 82{,}500$

The first firm's profit will increase by 6.9% while its competitor's profit will decrease by 9.17%! Is that a perfect way to compete? This means that both firms have an incentive to increase the supply when they are at or near the profit-maximizing supply quantities.

4. The reason for adopting NE

Let us do the similar exercise for NE. We know that there is no reason for a firm to decrease its supply. However, a firm may be interested in increasing supply. Consider that firm 1 increases its supply quantity from NE by 50 so $q_1^* = 450, q_2^* = 400$. The total supply now will be 850. The price will be \$575.

$\pi_1 = (575 - 400) * 450 = 78{,}750$

$\pi_2 = (575 - 400) * 400 = 70{,}000$

The profit for firm 1 reduces by 1.6%, while for firm 2, by 12.5%. Although it will hurt its opponent more, it reduces its own profit some, too. Therefore, the strategy is not rational.

The incentive to restrict the supply quantity at NE for high and robust profit leads to higher cost for the consumers, though.

5. Practical considerations

In practice, such as in the case of Coke vs. Pepsi, or Boeing vs. Airbus, or UPS vs FedEx, the price-supply function may not be available. Even if a firm can figure it out, it may not be linear. Their production costs may be complicated and can possibly be a function of the supply quantity. Their costs are unlikely the same. Their products are sure not considered identical by some customers; just ask an Atlanta resident or Chicagoan about Coke versus Pepsi.

Although there are a lot of inaccuracies, the model still provides great insight into the competition. In a duopoly, or even in a tight oligopoly, meaning just a few dominant suppliers, the increase of supply quantity of one supplier can lead to price reduction, and can lead to a price war that hurts all competitors.

4. COURNOT COMPETITION AND THE PRISONER'S DILEMMA IN GAME THEORY

It is interesting that Cournot competition is an application of a very useful game called the prison's dilemma in game theory. In game theory, there are different types of games. They can be cooperative or noncooperative games. In cooperative games, the players cooperate to get the most payoffs for all the players as a whole, such as between cooperating supply-chain partners. The cooperative games can also be useful in bargaining games. In the noncooperative games, players play to win, such as in competition. The games can be static or dynamic. In static games, each player plays once at the same time. In a supply chain, this can be determination of the supply quantity or price for one time, such as in the fashion industry or special events products. In the dynamic game, decisions are made in succession. Our in-class duopoly game was a dynamic game in which we can continued to play to improve our profits over repeated decisions simultaneously over time, such as determining the weekly supply quantity. The decisions can also be made in response to the other player's actions, such as in chess, or changing online prices after observing the prices of the others.

Cournot competition is a noncooperative game. In practice, it can be a static or dynamic game. The competitors determine supply quantities for one time or in each planning cycle.

In the Cournot game, the assumptions are that competitors have indistinguishable products with the same cost. Their strategy is to determine the supply quantity simultaneously. The combined supply quantity determines the profit based on the price set by a known reverse-demand curve. Their payoffs are the profits. The object is to maximize profit.

4.1 Prisoner's dilemma

The prisoner's dilemma game was created by Dresher and Flood in the 1950s. The story was that an armed robbery occurred. The police arrested two suspects. The police were convinced that they were guilty; however, there was insufficient evidence to convict them unless they confessed. The police interrogated them individually, and offered each suspect the following:

- Since you carried a gun when caught, this you will get one year in jail.
- If you confess but your partner does not, you go free and your partner gets ten years.
- If you both confess, each of you will get five years.

The suspects would like to find a way to get minimum jail time.

4.2 Game representation in strategic form

We will use the following symbols to define the game:

1. 2 players: A, B
2. $a_1, a_2, \ldots\ a_m$: Player A's m possible strategies
3. $b_1, b_2, \ldots\ b_n$: Player B's n possible strategies
4. $\pi_A(a_i, b_j)$: Player A's pay off given both strategies
5. $\pi_B(a_i, b_j)$: Player B's pay off give both strategies
6. Game structure: 2 players, static, simultaneous move.

There are many ways to represent games. We will use the normal or strategic form. The strategic form is presented in a bi-matrix, or a matrix with two entries in each cell, as shown in the table below. Another popular form is called the extensive form, in which the different decisions are represented by different branches in a tree structure.

$\boldsymbol{\pi_A \backslash \pi_B}$	$\boldsymbol{b_1}$	$\boldsymbol{b_2}$		$\boldsymbol{b_n}$
a_1	$\pi_A(a_1, b_1), \pi_B(a_1, b_1)$	$\pi_A(a_1, b_2), \pi_B(a_1, b_2)$		…
a_2	$\pi_A(a_2, b_1), \pi_B(a_2, b_1)$	…		…
a_m…				

The decision is kept secret from his opponent. The objective for each is to minimize jail time. We can use negative numbers to represent the jail times in the bi-matrix because they are undesirable.

$\pi_A \backslash \pi_B$	Silent	Testify
Silent	1, 1	0, 10
Testify	10, 0	5, 5

We will see below that a rational suspect would testify, because no matter what the other suspect selected, testify would get shorter jail time.

In general, if the payoffs have the following form in a bi-matrix, it is called the prisoner's dilemma.

$\boldsymbol{\pi_A \backslash \pi_B}$	$\boldsymbol{b_1}$	$\boldsymbol{b_2}$
a_1	B, B	D, A
a_2	A, D	C, C

Where A > B > C > D

The prisoner's dilemma has a wider variety of applications, such as the strategies of the two political parties in the United States, oil supply strategies by the OPEC and Russia, or CO2 emissions.

4.3 Cournot competition is an application of the prisoner's dilemma

In the duopoly Cournot competition example, the two players are 1 and 2. Each selects a supply quantity to compete. Based on the analysis above, each has plausible strategies between the maximum profit supply quantity of 300, or NE supply quantity of 400 each. We can enter the supply strategies and associated payoff (in thousands of dollars) for each in a bi-matrix shown in the table below:

$\pi_1 \backslash \pi_2$ in thousands $		q_2	
		300	400
q_1	300	90, 90	75, 100
	400	100, 75	80, 80

A = 100, B = 90, C = 80, and D = 75.

There are several possible strategy combinations between the winner and losers.

1. The winner supplies NE quantity of 400 with winning payoff of 100K. The loser supplies the maximum profit quantity of 300 with losing payoff of 75K.
2. Both suppliers supply NE quantity of 400 with equal pay of 80K.
3. Both suppliers supply maximum profit quantity of 300 with equal payoff of 90K.

The main difference between this decision and other decisions is that one player's payoff depends on the action of the other player. How do we use the strategic form to select the strategy?

4.4 The concept of dominance and dominated action

We assume that the players are rational, meaning they will select strategies with maximum payoff. Sometimes, the best strategy for a player is independent of the opponent's actions. If payoff of a strategy *i* for a player is worse than the payoff of another strategy *j* regardless of its opponent's strategy, the strategy *i* is strictly dominated by strategy *j* and therefore can be removed from consideration. A strategy can also be weakly dominated by a combination of other weighted strategies.

4.5 Finding Nash equilibrium based on elimination of dominated strategies

We can apply successive removal of dominated strategies to find the Nash equilibrium. Consider the same example. We can follow these steps and refer to the bi-matrix table below.

1. For player 1, supply 300 would lead to payoffs of 90K if the opponent supplies 300 or 75K if the opponent supplies 400. However, supplying 400 would lead to 100K and 80K, respectively, both higher than supplying 300. Therefore, $q_1 = 300$ is a dominated strategy. It can be removed or crossed out as shown.

2. For player 2, the logic is identical. Therefore, $q_2 = 300$ is a dominated strategy. It can be removed.
3. After removal, the only remaining strategy is therefore $q_1 = q_2 = 400$, or NE supply quantity. The profit will be 80 each, instead of 90 each at maximum.

$\pi_1 \backslash \pi_2$ in thousands $		q_2	
		300	400
q_1	~~300~~	~~90, 90~~	~~75, 100~~
	400	100, 75	**80, 80**

This is the same as applying first-order conditions.

5. COURNOT COMPETITION EXTENSIONS

You can extend Cournot competition to include multiple players, to have different costs, and so on.

5.1 *n*-players

In most markets, there are more suppliers than two to compete with. For example, there are more than 10 major car suppliers, 20 major third-party logistics companies (3PLs), many electronics suppliers, hundreds of fashion stores, and countless dry cleaners and travel agents. You can find the NE in the similar fashion as duopoly: price-supply curve, make the assumptions about the same cost, common price, undifferentiable product, apply the first-order condition to solve multiple simultaneous equations, then assume they all supply the identical quantities to find the supply quantities, prices, and profits at NEs. The formulas for the solutions are summarized in Table 4.1. The trend is shown in Figure 4.5.

Table 4.1 Formulas for the monopoly, duopoly and n suppliers.

	Quantity	**Price**	**Individual Profit**	**Total profit**
Monopoly	$\frac{P-c}{2B}$	$\frac{P+c}{2}$	$\frac{(P-c)^2}{4B}$	$\frac{(P-c)^2}{4B}$
Duopoly, Cournot	$\frac{P-c}{3B}$	$\frac{P+2c}{3}$	$\frac{(P-c)^2}{9B}$	$\frac{2(P-c)^2}{9B}$
Oligopoly with *n* suppliers	$\frac{P-c}{(n+1)B}$	$\frac{P+nc}{(n+1)}$	$\frac{(P-c)^2}{(n+1)^2B}$	$\frac{n(P-c)^2}{(n+1)^2B}$

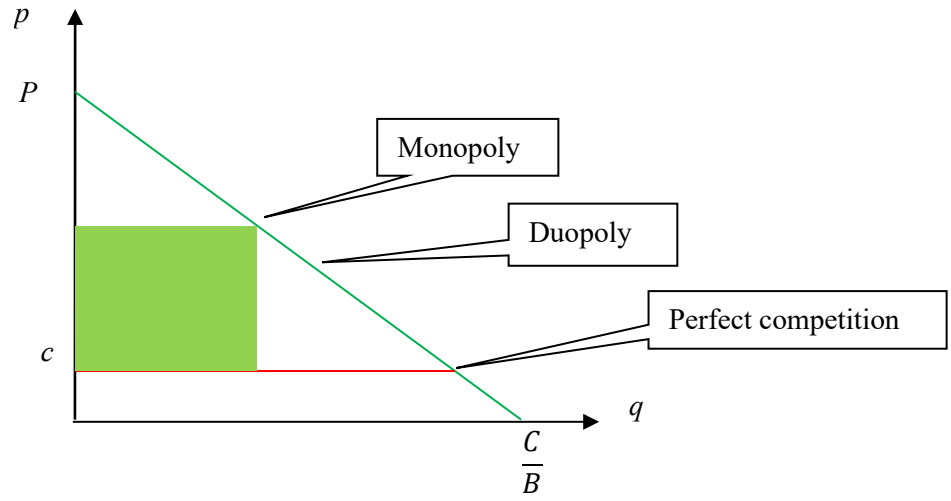


Figure 4.5. The total supply quantity and price from monopoly to n-player Cournot competition.

This chart also shows a firm's incentive to grow, to merge, and to acquire. Most business evolved from cottage industry to the fragmented and then loose oligopoly, then to tight oligopoly. Based on Cournot and Figure 4.5, fewer suppliers can lead to benefits in both higher prices per unit and more supply from each firm. One way to achieve this is through mergers and acquisitions.

5.2 Duopoly with cost reduction

Firms always strive to reduce costs. What is the impact of cost reduction to the supply quantity, resulting price, and the profits? There are two different situations depending on the information transparency or privacy.

5.2.1 Firm 1 reduces the cost and keeps the information private

If one of the players lowered its cost to c_1, we can follow the similar derivation to find the NE with the new information. If firm 2 does not know firm 1 has reduced its price, it would assume the same cost. The profit functions for each player with different costs are:

$$\pi_1(q_1, q_2) = q_1[P - c_1 - B(q_1 + q_2)]$$

$$\pi_2(q_1, q_2) = q_2[P - c - B(q_1 + q_2)]$$

Apply FOC, we have the

$$q_1^* = \frac{P - c_1}{2B} - \frac{q_2}{2} \quad (5)$$

$$q_2^* = \frac{P-c}{2B} - \frac{q_1}{2} \quad (6)$$

Based on the assumption that firm 1 kept the cost reduction private, firm 2 has no reason to change it supply quantity, or

$$q_2^N = \frac{P-c}{3B}$$

Firm 1 can use this quantity to find its own new NE supply quantity, or

$$q_1^N = \frac{P-c_1}{2B} - \frac{q_2^N}{2} = \frac{P-c_1}{2B} - \frac{P-c}{2*3B} = \frac{2P-3c_1+c}{6B} = \frac{P-c_1}{3B} + \frac{c-c_1}{6B}$$

Here

$$\frac{P-c_1}{3B} > \frac{P-c}{3B}$$

$$\frac{c-c_1}{6B} > 0$$

Therefore, $q_1^N > q_2^N$. Since q_2^N stays the same as without cost reduction but q_1^N become higher, the total quantity will be higher. The price will be lower, or $p(c_1, c) < p(c,\ c)$. This will lead to higher profit for firm 1 and lower profit for firm 2.

5.2.2 When firm 1 reduces its cost and firm 2 finds out

If you are the supply-chain engineer or manager in firm 2, how would you respond if you noticed the increased supply from firm 1 and the price drop it caused? Based on the modeling assumptions, the increased supply from firm 1 will reduce its own profit if the cost stays the same. Firms strive to achieve lower costs all the time. You may guess that firm 1 lowered their cost. Assume that through business intelligence, you figured out that the firm 1 has reduced the cost to c_1, how much should you supply? You can develop the new NE by solving simultaneous equations

$$q_1^N = \frac{P-c_1}{2B} - \frac{q_2^N}{2}$$

$$q_2^N = \frac{P-c}{2B} - \frac{q_1^N}{2}$$

The solution will be

$$q_1^N = \frac{P - c_1 + (c - c_1)}{3B} = \frac{P - c_1}{3B} + \frac{c - c_1}{3B}$$

$$q_2^N = \frac{P - c + (c_1 - c)}{3B} = \frac{P - c}{3B} - \frac{c - c_1}{3B}$$

You will find that firm 1 is supplying its NE quantity based on the new cost. You will also find that the new NE quantity for firm 2 is lower than previous NE quantity. The total combined quantity will be lower, so the price will go higher. The profit for firm 1 will be even higher while the profit for firm 2 will be lower. This shows the importance of lowering cost in competition!

Example

Consider the same paper example. Firm 1 reduced its marginal cost significantly from 400 to 300, a 25% reduction. Please find the supply quantities, the resulting price, and the profits if

1. Firm 2 does not know firm 1 has reduced the cost.
2. Firm 2 found that firm 1 reduced its cost to 300 (this is a big if).
3. Discussions

Solutions

1. Firm 2 does not know firm 1 has reduced the cost.

The supply quantity for firm 2 stays at $q_2^N = 400$. The supply quantity for firm 1 is

$$q_1^N = \frac{P-c_1}{2B} - \frac{q_2}{2} = \frac{1000-300}{2*0.5} - \frac{400}{2} = 500$$

$$p = 1000 - 0.5(500 + 400) = 550$$

$$\pi_1 = (550 - 300) * 500 = 12{,}500$$

$$\pi_2 = (550 - 400) * 400 = 60{,}000$$

2. Firm 2 does not know firm 1 has reduced the cost.

$$q_1^N = \frac{P-c_1}{3B} + \frac{c-c_1}{3B} = \frac{1000-300}{3*0.5} + \frac{400-300}{3*0.5} = 533.33 = 533$$

$$q_2^N = \frac{P-c}{3B} - \frac{c-c_1}{3B} = 400 - 66.66 = 333.33 = 333$$

The total quantity reduced from 900 to 866. The price should go higher

$$p = 1000 - 0.5(533 + 333) = 567$$

$$\pi_1 = (567 - 300) * 533 = 142{,}311$$

$$\pi_2 = (567 - 400) * 333 = 55{,}611$$

3. Firm 1's profit is even higher while firm 2's is lower. This is counter intuitive. The result is based on a model with many assumptions. It suggests that firm 2 voluntarily reducing production. In practice, these assumptions may not hold.

5.3 Importance of information

The theoretical analysis and results are under very strict assumptions. In practice, there are many other considerations to be discussed later. Do not apply these results literally to all problems. You should learn the modeling: assumptions, derivations, results, and symbolic analysis.

5.4 Tacit Collusion

In the game, you may find that the Nash equilibrium solution in the Cournot model is not the profit-maximizing solution. You may have noticed that if both players produce half of the monopoly quantity, the price and profit for each firm will be higher than that at the NE. However, the only way to get there is through collusion. However, collusion is illegal. You can find many cases of allegation, indictment, and convictions of collusion or price fixing on the internet. These firms were accused in collusion through tacit signals or activities without public record .

Supply-chain engineers or managers who had studied the Cournot model would understand its benefits. The profit-maximizing objective incentivizes tacit collusion. However, this strategy avoids competition, a fundamental force for progress, and hurts the clients, and that is why it is illegal.

6. BERTRAND COMPETITION

Bertrand reviewed the Cournot competition and concluded that Cournot's assumption was wrong about the decisions suppliers made to compete. He believed that the suppliers decide on the prices to compete in the marketplace instead of supply quantity.

6.1 Modeling

Assumptions from the Cournot model

1. Firms 1 and 2 supply identical products in the same market.
2. The supply costs per unit are the same and transparent for both firms.

Additional assumptions

3. The supplier with lower price gets the entire market
4. The resulting price is determined by the supply quantity of the lower price
5. If their prices are the same, the two suppliers share the market equally

Decision: the firms determine the price to maximize profit.

Mathematically, we can express the quantity sold by firm 1 below

$$q_1(p_1, p_2) = \begin{cases} 0 & \text{if } p_1 > p_2 \\ 0.5q_1(p_1) & \text{if } p_1 = p_2 \\ q_1(p_1) & \text{if } p_1 < p_2 \end{cases}$$

Let p_{low} be the lower or winning price; we then can use the demand model,

$$q = Q - B'p_{low}$$

where *B'* is a slope, which is the inverse of *B* in the reverse demand model $p = P - B'q$, or *B*' = B^{-1}. If you feel more comfortable with the original price-quantity relationship, it will work also. The profit for the firm setting a lower price is

$$\pi_{low}(p_{low}, p_{high}) = \begin{cases} (p_{low} - c)(Q - B'p_{low}) & if\ p_{low} < p_{high} \\ 0.5(p_{low} - c)(Q - B'p_{low}) & if\ p_{low} = p_{high} \end{cases}$$

Example

Two firms are to supply very similar products for the holiday season. The cost of supply one container is $1 million. The firm predetermines its prices, and produces half of the product ready to sell, based on prior record, the maximum possible price is $5 million per container. The maximum supply quantity is five containers. The two firms can set the price once. If firm 1 sets the price lower, it will capture the entire market. If the prices are identical, they split. If the firm's price is higher, it will incur $0.2 million to get ready to supply the product.

Assumptions

1. The demand curve is accurate
2. Simple Bertrand game where the lower-price supplier takes all the market, and equal prices share the market
3. Assuming the unsold products incur cost but are of no value.

We can find $P = 5$, $B = 1$, $c = 1$. The demand curve or demand model is

$$q = 5 - p$$

Let us try some scenarios.

If $p_1 = 2$, and $p_2 = 4$. Firm 1 takes all market.

Then, $q_1 = 3, \pi_1 = (p_1 - c)q = (2-1) * (5-2) = 1 * 3 = 3, \pi_2 = 0.$

If $p_1 = p_2 = 2$. Two firms will share the market. However, each can only sell half of what they supplied.

$$q_1 = q_2 = q = 1.5,\ \pi_1 = \pi_2 = (p-c)q = (2-1) * 1.5 = 1.5.$$

Note that if your price is higher there is no revenue but there can be cost in the planning and preparation to set up the product to sell.

The payoffs in normal form in a bi-matrix presentation in game theory is show below. The first headers in the columns and rows are the prices by each firm. The first entry in each cell is the profit for firm 1. The second entry in the cell is the profit for firm 2. The boldface entry is the NE solution at (1, 1). The cells with underlined values are the maximum and equal profit by both firms. However, it is unstable.

Firm 1 \ Firm 2	**1**	**2**	**3**	**4**
1	0, 0	0, 0	0, 0	0, 0
2	0, 0	1.5, 1.5	3, 0	3, 0
3	0, 0	0, 3	**2, 2**	4, 0
4	0, 0	0, 3	0, 4	1,5, 1,5

6.2 Playing the Bertrand game

Let us play a game using the parameters in the special paper supply monopoly example. Recall the first term in (1) and (2) are the profit maximizing quantity for monopoly, which is 6000. Anything above 6000 will lead to loss with just one player. Therefore, it is not good idea. In fact, you want to be significantly less than 6000 because the other supplier also supplies.

We can use TurningPoint or MobLab applications to play this game. Make sure you can enter. The objective can be maximum profit or beating the opponent. I will let you determine your objectives.

The preparations

1. Form a pair to compete. Remember your competition's name.
2. Find a piece of paper or start a spreadsheet shown below.
3. Create a table on paper or computer with the following columns.
4. Record the date, your name, and your competition's name.
5. Review the monopoly optimum quantity, equation (1) and (2)

The game record form

Round	Your p	Unit profit p - c	$q = P - Bp$	Your π
1				
2				
…				
Sum				

6.3 Nash equilibrium

At any price higher than opponents, a firm risk cost without revenue. Each firm has incentive to lower the price to increase get some profit and avoid cost. The competition will lead to price war. The only stable strategy for both, or *NE*, is when $p_1 = p_2 = c$. At cost is not fatal because all the costs are covered, including compensations. This is very different from the results in Cournot competition. You can see the importance of assumptions!

In reality, products are never identical. Competitors do not have simply share the market even if they set the same price at cost. The firms may adopt a cost-plus strategy to set some margin. The main difference between Cournot and Bertrand Duopoly is that if the competitors engage in the former, the competitors can make substantial profits. On the other hand, if the competitors engage in a price war, they will both reduce their profit. Price transparency tends to push for the price wars and lead to lower profit margins for the competitors.

6.4 Bertrand extension on different costs

The above conclusions are based on when the costs are the same. Firms are in constant search to lower the cost. If you can reduce your cost, what should be the best strategy?

If firm 1 reduces the cost while firm 2 has the same cost, or we have $c_1 < c_2$. If you set the price at any level $c_1 < p_1 < c_2$, you will make a profit of $p_1 - c_1$ for each product and get the whole market of $P - Bp_1$. Firm 2 cannot compete, because it would lose money at price below their cost. That is why cutting cost is so important. The profit for firm 1 will be

$$\pi_1(p_1) = (p_1 - c_1)q_1 = (p_1 - c_1)(P - Bp_1)$$

You will maximize this by setting the price as close as c_2 as possible, but low enough to prevent firm 2 from getting market share.

7. COMPETITION WITH DIFFERENTIATION

What have you learned from simple Bertrand competition and Cournot competition with identical products? Price wars on identical products or services leads to lower prices for the consumers at the price of lower profit for the firms. Note that these models assume identical products. Therefore, the firms can also compete through differentiation. For example, after Pepsi came with Pepsi 1, Coke came out with Coke 0. One express service provider provides overnight service, another guarantee 8 a.m. delivery. iPhone has been successfully differentiated from other smartphones for many years. Once the products or services are differentiated, firms can gain some

monopoly power. When a firm increases its price, the demands for its product may not decrease because the competitor's products are different. On the other hand, if the competitor increases its price, the demand for the firm can increase.

7.1 Dimensions to differ

In a supply chain, a supplier or a service provider can differentiate from the competition in many dimensions. We will discuss many of these later in detail.

1. Speed, such as lead times from time of order to the time of shipping
2. Service levels: the percentage of goods or services delivered as planned
3. Information connectivity: EDI, supply-chain software
4. Consistency in product and service (quality, on-time, PPM)
5. Narrow delivery window
6. Flexibility to volumes, terms, change requests
7. Variety of products or services: customers may desire one-stop shopping
8. Returns, warranty
9. Terms
10. Shipping charges
11. Brand, trust, taste, accountability, relationship
12. Perceived quality
13. Appearance
14. Individual cost to acquire (more driving, more paperwork)
15. Labeling (mattresses)
16. Others.

The more your product or service differs from your competitor's, the more "monopoly" power you can gain.

7.2 Cost and benefit to differentiate

Differentiation may come with a cost. Figure 4.6 shows the relationship between the cost incurred to achieve the difference and the perceived positive difference. The lower the cost to differentiate and higher the positive perceived differences are better.

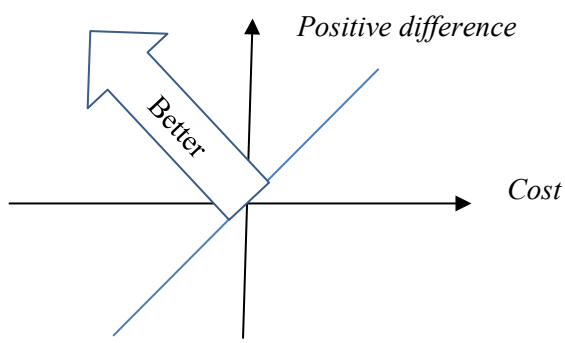


Figure 4.6 Cost and the resulting positive difference differentiation.

What one wants to achieve is the largest positive differentiation with minimum cost, or least negative differentiation with largest cost savings. The best place is in the second quadrant, when you can save money and gain positive difference. Coke came up with a 250-mL slim can in July 2013. It cost less to bottle, ship (except extra set up cost!), and so on, and they charged more. This required out of the box thinking and generated game-changing product or services. Another example is detergent refill. The simpler packages cost less. The initial advertisement only quoted the lower cost due to packaging (also transportation due to tighter packs). However, it was the added message of environmental issues that really resonated with many. A company can even profit more from the sales.

Note: whenever dealing with environmental issues, it is often more involved. When Walmart worked with Tide to increase the detergent with higher concentration, it claimed reduction in CO_2 emissions due to reduced transportation energy involved. It was good. It also claimed the reduction of water. Is this true? In addition, the higher-concentration detergent may not dissolve as well as the low-concentration detergent. What is the environmental impact of undissolved chemicals in the drainage system?

7.3 The law of minimum differentiation in location

Consider two suppliers with identical products. Both want to find the best location to maximize the profit.

Assumptions

1. Two suppliers supply identical products with identical costs and price.
2. They can select location along a one-dimensional line.
3. The customers are distributed uniformly along the line.
4. The customer would visit the store that is closer.
5. They want to find the store location to maximize their share of customers (to maximize profit, since both have the same price and cost).

Example

Two mobile stores sell identical goods along the main street, which is four miles long. They have the same cost and sell at the same price. The customers go to the closest store to get the goods. There are two customers in each mile, uniformly distributed.



Where would you locate your store?

If store 1 moves one mile toward the center, what will be the impact?



In this case, all customers to the left of the vertical line, or five of them, will buy from store 1. The remaining three will buy from store 2. Store 2 would want to do the same.

7.3.1 Clicker or MobLab game

What can we conclude in this game?

1. The player closer to the center get more customers.
2. If the locations are symmetric, the stores share the market equally.
3. The stable solution is at the center with bold face when both stores get equal share of the customers.

1 \ 2	0	1	2	3	4
0	4, 4	1, 7	2, 6	3, 5	4, 4
1		4, 4	3, 5	4, 4	5, 3
2			4, 4	5, 3	6, 2
3				4, 4	7, 1
4					4, 4

7.3.2 Equilibrium

Consider supplier 1 at position 0, which can increase its share of customers by moving toward the center. This is the same for supplier 2. Therefore, both would locate at the center to share 50% of the customer each.

Why two opposite laws? The law of minimum differentiation applies when two firms provide identical products. This can be selection of the physical location, the lead time, the PPM, and so on. If a firm wants to differ, the maximum differentiation applies—differentiating the most that is feasible or realizable.

From a service perspective, the stores provide better service to customers by locating at one-mile and three-mile markers, where the customers only have to travel between zero and one miles. If they are both in the center, the travel would be between zero and two miles. Locating in one and three can make sense if two stores belong to the same company; the manager's rewards are by the total volume. However, such a reward system can lead to moral hazard: not to keep the store clean, bad service, and so on. It can reduce the total demand.

7.4 The Bertrand competition with differentiated price

The firms work hard to make sure the products or services are not identical. They can design them to differ in many dimensions. As a result, they command different prices. Different prices may impact the demand of each competitor differently. This can be modeled as supply quantity as function of the prices of both at different levels:

$$q_1 = Q_1 - B_{11}p_1 + B_{12}p_2$$

$$q_2 = Q_2 - B_{22}p_2 + B_{21}p_1$$

Where *Qs* and B_{ij} are the positive constants. Q_1 is the maximum quantity that firm 1 expects to sell. q_1 decreases with the p_1 at a constant rate of B_{11}. In a competitive market, the increase of price by one firm moves customer to its competitor with similar products. That is why this term is negative. Conversely, if the competitor raises its price, it can increase its volume; that is why q_1 increases with the p_2 at a constant rate of B_{12}. Please see if you can explain in the demand change for firm 2.

7.4.1 Profit and best response functions

The profit functions for the firms are:

$$\pi_1(p_1, p_2) = (p_1 - c_1)q_1 = (p_1 - c_1)(Q_1 - B_{11}p_1 + B_{12}p_2)$$

$$\pi_2(p_1, p_2) = (p_2 - c_2)q_1 = (p_2 - c_2)(Q_2 - B_{22}p_1 + B_{21}p_2)$$

Apply first order conditions, we can find the best response functions for the firms

$$p_1^* = \left(\frac{Q_1}{2B_{11}} + \frac{c_1}{2}\right) + \frac{B_{12}}{2B_{11}}p_2$$

$$p_2^* = \left(\frac{Q_2}{2B_{22}} + \frac{c_2}{2}\right) + \frac{B_{21}}{2B_{22}}p_1$$

The parameters in the first term for each firm depends on its own set of parameters. It increases with the constant supply quantity and cost, and decreases with the slope. The second term is proportional to the ratio of the rate of increase and the rate of decrease.

The major difference between the Bertrand price competition with differentiation and Cournot model is the sign of the second term. Here, the best response price increases with the opponent's prices while in Cournot, the best response supply quantity decreases with the opponent's supply quantity. This change can create win-win situation for both firms.

The larger the constants B_{ij}s, the more sensitive the market size to the supplier. The first term is for the influence from its own prices. The second term is from the competition.

Example 1

Consider the special paper supply in the Cournot model. The two firms are forced to compete on price, which leads to lower and lower prices. Both firms invest to make their product different so that they can stay away from cutthroat price competition. Firm 1 modified its paper product to make it work better with colors and better feel. Firm 2 modified its supply chain to provide faster and better services. The monthly equivalent investment cost is $10,000 per month for both firms.

The marginal cost increased $20 to become $420 for both firms. Their supply quantity as function of the prices can be modeled as

$$q_1 = 400 - 2p_1 + 2p_2$$

$$q_2 = 400 - 2p_2 + 3p_1$$

In the Cournot competition, 400 was the NE. Their investment in differentiation can either increase the supply when the competition increases the price or decrease due to its own price increase. For firm 1, the impact is symmetric. For firm 2, the increase in price by firm 1 will benefit it more, or customers prefer faster service to easier printing.

Solutions

The profit for firm 1 is

$$\pi_1(p_1, p_2) = (p_1 - c)q_1 = (p_1 - 420)(400 - 2p_1 + 2p_2)$$

Applying the first-order condition, we have

$$\frac{\partial \pi_1}{\partial p_1} = 1 * (400 - 2p_1 + 2p_2) + (p_1 - 420)(-2) = 1240 - 4p_1 + 2p_2 \Rightarrow 0$$

The best response function for firm 1 is

$$p_1 = 310 + 0.5p_2$$

Similarly,

$$\pi_2(p_1, p_2) = (p_2 - c)q_2 = (p_2 - 420)(400 - 2p_2 + 3p_1)$$

$$\frac{\partial \pi_2}{\partial p_2} = 1 * (400 - 2p_2 + 3p_1) + (p_2 - 420)(-2) = 1240 - 4p_2 + 3p_1 \Longrightarrow 0$$

The best response function for firm 2 is

$$p_2 = 310 + .75p_1$$

Note the major difference between this response function and Cournot response function is that the minus sign changed to a plus sign. The opponent's price increase led to higher prices of its own!

At NE, we solve simultaneous equations,

$$p_1 = 310 + 0.5(310 + .75p_1) = 465 + 0.375p_1$$

$$p_1 = 744$$

$$p_2 = 310 + .75 * 744 = 868$$

These prices were much higher than the Cournot commodity NE prices of \$600. The quantity sold would be

$$q_1 = 400 - 2p_1 + 2p_2 = 400 - 2 * 744 + 2 * 868 = 648$$

$$q_2 = 400 - 2p_2 + 3p_1 = 400 - 2 * 868 + 3 * 744 = 896$$

The profits

$$\pi_1 = (p_1 - c)q_1 - 10000 = (744 - 420) * 648 - 10000 = 199{,}952$$

$$\pi_2 = (868 - 420) * 896 - 10{,}000 = 391{,}408$$

The investment into differentiation benefited both firms with larger market and higher profit per unit for both firms. However, these results are based on the specific parameters we set. The change of parameters can lead to higher or lower levels of increase. The profit can also decrease in certain conditions. This example only intends to provide an overview.

8. SUMMARY OF ECONOMIES OF COMPETITION

The Cournot and Bertrand models provide insights for the firms to formulate strategies for effective competition. In tight oligopoly or a few suppliers supply very similar products or services with the similar cost, they should control the supply quantity captured in the Cournot model. First, they should avoid a price war explained in the Bertrand model.

Table 4.2 Comparison of Cournot and Bertrand competition.

Game	Decision variable	Best response	Decision value for highest profit	Decision value at NE	Profit at NE
Cournot	q	$q_1^* = \frac{P-c}{2B} - \frac{q_2}{2}$	$\frac{P-c}{4B}$	$\frac{P-c}{3B}$	$\frac{(P-c)^2}{9B}$
Bertrand	p	$p_1^N = c$	$\frac{P+c}{2}$	c	0
Bertrand differentiation	p	$p_1^* = \left(\frac{Q_1}{2B_{11}} + \frac{c_1}{2}\right) + \frac{B_{12}}{2B_{11}} p_2$	Complex	Complex	Complex

The second effective way in this situation is to increase profit through cost reduction. The benefits of cost reduction are captured in both the Cournot and Bertrand models. The third effective way to achieve higher profit is through product differentiation, which allows a firm to enjoy some monopoly power. The profit increase through differentiation can be modeled by the Bertrand price competition with differentiated products. The parameters and best response curves are summarized in Table 4.2. The main difference between Cournot and Bertrand differentiation models are the "-" vs "+" sign in the best response function, colored in red!

In a loose oligopoly, many suppliers of commodities market are price takers. They are not in position to control their profit through volume or price. The best way to increase profit is through cost reduction. They can also differentiate their product or services to avoid cutthroat competition and enjoy some monopoly power. Since the profit decreases with the number of suppliers, they can also increase profit by reducing the competitors through growth, or more efficiently, through mergers and acquisitions.

The most profitable situation is a monopoly. The way to reach monopoly to develop totally new products or services. Netflix dethroned Blockbusters; cellphones replaced many things we used to use, such as phones, calendars, cameras, watches, and note pads. Airbnb challenges the hotel industry, and Amazon challenges brick-and-mortar and fixed price medical services compete with conventional opaque price services.

8.1 Practical considerations

Models are based on assumptions. On the ground, very few products are truly identical. Maybe no. 2 corn or no. 2 pencils. The price-supply curve is unknown or nonlinear. Let us first consider Boeing and Airbus. They are true duopoly. There have not been other credible suppliers for a while for large commercial jets. Boeing 737 and Airbus 320 are very similar products in size and flying characteristics. They all try to improve and to differentiate all the time. They also enjoy various government programs. Therefore, the Cournot and Bertrand models can only be used for insights, not for actual supply quantity or price calculations. Let us consider Coke and Pepsi. In the beverage market, they are much bigger than many much smaller suppliers. Both Cournot and

Bertrand provide insights. The two large firms enjoy higher prices and higher profit margins than commodity beverage firms do. A similar analogy applies to UPS and FedEx.

The internet made prices more transparent and easily compared. The price transparency makes the Bertrand model more prevalent. It is better for the client but reduces profit for the supplier. The suppliers often work to make the prices complicated and more difficult to compare. For example, cars have MSRP, medical drugs publish list prices, and freight carriers let anyone find the prices by entering the shipping information through an open form. This is public information to let the client compare and select cost leaders. In practice, the final prices are negotiated and become opaque to the public. The clients are happy with the discount. The supplier knows their cost and can use the Cournot model to maximize their profit. This scheme benefits the supplier overall. Imagine if the actual prices for cars, drugs, or shipping rates were public! This would make the Bertrand model more relevant. It would benefit the client and cut profit from the supplier. There is already an effort through the internet to publish the actual prices for drugs.

8.2. Strategies in a supply chain in different organizational structures

A supply-chain engineer or manager may work for a supplier that is a monopoly, one of the firms in a tight oligopoly, or one firm in loose oligopoly. You may also work in a firm that is a client to a supplier that is a monopoly, tight oligopoly, or loose oligopoly. How do you apply the insights from the Cournot and Bertrand models to formulate your strategies? Table 4.3 is a summary of strategies derived from the insights from these two models.

Table 4.3. Supply-chain strategy based on different markets as a supplier or a client.

	As a supplier and as a	**As a client of a**
Monopoly	Watch out for substitute (Blockbuster Video, PDAs, small digital cameras, data storage equipment)	Find alternatives Threat with anti-trust laws
Tight oligopoly	Avoid price wars Avoid price transparency Reduce cost Differentiate	Encourage price wars Work on price transparency Watch out for tacit collusions
Loose oligopoly	Differentiate Reduce cost Mergers and acquisition	Make supplier compete Watch out for fake differentiation

9. COMPETITION AND HUMAN NEEDS

What are the impacts of competition to supply for human needs? Today, people in the developed economy have enjoyed abundance of supply of goods and services for basic needs at low prices

for decades. Competition in the markets incentivizes suppliers to reduce cost to improve efficiency, to improve products and services, and to create new products and services. The competition among information technology firms increases the transparency and information availability. The markets have done their work! Throughout this book, we will discuss strategies that help a firm to succeed in competition.

However, not all competition strategies naturally lead to better supplying human needs.

9.1 Strategies good for profit but that may not be good for human needs

Not all strategies in competition lead to better supplying human needs. For the products and services that are important for human needs, especially the ones important for a low-income population, here are some examples:

1. Collude with the competitors tacitly. The collusion would lead to higher prices.
2. Absorb the competition through mergers and acquisition to avoid competition. In the Porter's model for competition and Cournot extension, we can see that the reduction in the number of suppliers can lead to higher prices. You may have noticed that hospital mergers take place at a very fast pace. Higher prices are especially true if the firm being merged incurs a lower cost in supply.
3. Differentiate through labeling or marketing the essentially identical products or services through different channels. This strategy leads to higher prices without better service to human needs.
4. Differentiate to lure people's wants. There are many examples. In food, some supply with excessive sweetness or caffeine. In fashion, pre-wash or pre-damage the product to differentiate.
5. Take advantage of the established market domination to suffocate the new entrance of better products or services. Porter's model lists this strategy specifically.
6. Design the product to underperform when planned new products are launched. Apple and others have been accused of using this strategy.
7. Lock in the customer base through different accessories (such as chargers), phone numbers (used to be associated with a service provider), and locked phones. Such effort increases waste and reduces competition. Cellphone services used this strategy for many years. They still enjoy part of this strategy through contracts.

These types of competition do not always help to supply human needs. In fact, they can sometimes do the opposite.

9.2 The incentive to supply to over-consumption and human wants

As society become more affluent, more and more people have more and more disposable income. They can spend on what they want but they do not need, such as excess foods, the newest gadgets, rarely used high fashions, smoking, drugs, unhealthy indulgence, or over-cleaning. These goods are easier to differentiate and can command high profit margins. Therefore, the firms have incentive to supply to such wants. This by itself is not a major problem, and can potentially increase employment. However, they can also hurt certain human needs for themselves or others.

9.3 The challenges when human needs are external to the financial systems

Fresh air, arguably the most important human need, is mostly external to the financial systems. During the pandemic, there have been arguments about the use of masks. The problem is similar to fresh air. The decision is private. However, the impact is public, because an infected person without a mask can spread the virus through air, which can infect other people. It is similar with for vaccination. Vaccination can constrain the multiplication of virus in vaccinated people who are exposed to virus, and reduce the amount of virus the exposed person can spread into the air. Social distancing is also related to public virus spread. However, social distancing is in the financial system: it can reduce or close financial activities!

Although clean water is in the financial system, its financial value is not commensurate to its value to human needs. In game theory, the game "tragedy of the commons" describes a situation where the consumption of common goods can lead to the depletion of the goods. Other important public goods that are essential to human needs include fresh water, fish in the open waters, and a quiet space to relax or study.

9.4. Markets and human needs

Competition is good. Competition can align the interests of different firms to improve efficiency, lower cost, and create better products and service to supply for human needs. It is competition in the free market that brought abundance to society.

However, competition can also generate side effects when the pricing signals does depart from human needs, such as in food deserts, which hurt human nutritional needs. In addition, some competitions are not based on a true free market as it is intended, such as the healthcare supply chain. Too many players with conflicting interests are not always lined up with the patient's interests.

As the supply capacity reaches abundance in many markets, the supply chain starts to challenge the earth's carrying capacity. However, public goods, such as fresh air and fresh water, are not priced in relationship to their importance to human needs. Therefore, there is a need to consider the market design. Corporations have been more and more aware of the need to supply human needs. This will lead to major changes toward a better supply chain for human needs. Supply-chain professionals should check the firm's mission statement to ensure their strategies are aligned with their firm's mission, increasingly with social and sustainability dimensions—both are important human needs.

5. PRICING IN SUPPLY CHAIN

In Chapter 4 Competition, there are models related to pricing, such as Cournot, Bertrand, and Bertrand price differentiation. These models provide insights when the markets influence the prices through the demand and reverse demand models. They are useful with simple products or services that customers can compare and choose.

In the supply chain, the products or services between partners are often dissimilar and complicated. It is difficult for the client to compare the offers from different suppliers. In Chapters 2 Collaboration and Chapter 3 Principal-Agent Model can be used for the collaborating partners to set and accept or reject prices.

In certain markets, firms set their prices based on their costs or on their clients' private values they estimate. This leads to cost-plus pricing or auctions.

One aspect of revenue management relevant to pricing is to set prices differently for different clients to increase revenue on the same product or service. This way of price differentiation is also referred to as price discrimination, and can be based on age, credit score, quantity, memberships, socioeconomic status, or region.

Another way firms can compete or increase revenue is by setting prices dynamically over time. E-commerce has made dynamic pricing more convenient and it becomes more prevalent.

In this chapter, we will focus on cost-plus pricing, auctions, price discrimination, and dynamic pricing.

1. PRICING IN SUPPLY CHAIN

This is an overview of cost-plus pricing, auctions, price discrimination, and dynamic pricing.

1.1 Cost-plus pricing

Some firm set their prices based on the cost of supply plus some preset margin. As long as the firm can determine the cost of supply, they can add some percentage of the cost. This way, they can enjoy a known margin. Cost-plus pricing applies to non-commodity types of product or services when it is difficult to find alternatives, or it is difficult for the client to compare the products or services from different suppliers.

Determining the exact cost of a supply requires detailed record on materials, resource uses, utilities, times by engineers, managers, accountants, and direct labor. Keeping records can be demanding without automation. You can try to keep track of how you spend your own time in a day to experience the difficulty. The record changes with the variety of the products or services. On top of the above challenges, the values may experience randomness. Some statistical methods must be used to find the best parameters.

In practice, many firms may only keep track of the direct personnel time on the most important element. They lump the other complex costs into "overhead." This can lead to significant inaccuracies. Some supply requires more direct personnel time than others.

To improve accuracy, some firms use activity-based costing (ABC) to find more detailed costs associated with each product or service. The modern technologies in tracking and automated data

gathering can help to make ABC easier. The accuracy of ABC depends on the design of the system.

The drawback for the supplier is that the strategy is internally focused. There is no incentive to reduce the cost or improve quality or service because the cost will be recovered from the price. A supplier may not realize that it has become not competitive due to old technologies, inefficiencies, or poor incentives. It may be surprised one day to be outbid by other suppliers.

The drawback for the client is that cost-plus pricing encourages moral hazard. The supplier can pad up the costs to increase the price and profit! Car services, and many government and defense contractors, use cost-plus pricing. The cost of car repair at a dealership is typically set by labor hours. The hourly rate can be $50 per hour. The mechanic will only get a fraction of $50. The rest is overhead. They often spent much less time to finish the job than number of hours billed. Of course, there are variabilities. However, the dealership has a strong incentive to set the standard time longer than the average. In addition, their hourly rate is often much higher than those in independent or chain garages. With the transparency in car pricing, the dealership's margin on new cars is getting lower and lower. Service is a major part of their revenue.

1.2 Auctions

There are also business transactions that can be well described and then let the potential sellers or buyers bid. For example, an open lot can be useful for potential distribution centers or transportation hubs and can be auctioned off. Another example is for the inbound transportation service needs to be filled by carriers that a manufacturer wants to buy. These well-defined product or service needs can be prepared in the form of a call for proposals (CFP). The potential buyers of the lot, or the potential seller of services, can study the CFP and bid with prices or charges. The firm can select the winners among the bidders based on predefined rules. This is a form of an auction, an important mechanism to set the prices between supply chain firms.

1.3 Price discrimination

Firms can set prices differently on the same products or services based on the verifiable client characteristics or by customer's choice. This is one type of revenue management. Price differentiation in this manner economists call price discrimination.

Price discrimination can be based on verifiable client characteristics such as age, membership, consumption history, socioeconomic status, gender, institution, region, country, and so on. For example, student discounts, senior discounts membership discounts, frequent flyer benefits, supplemental nutrition assistance program (SNAP), commonly known as food stamps, government discounts, and different medical insurance plans. This type is referred to as direct price discrimination in revenue management.

Price discrimination can also be based on consumer preferences, such as many forms of quantity discounts, time of purchase in hotels and flights, and flexibility of specific product or service allocation, such as priceline.com. This type of price discrimination is referred to as indirect price discrimination.

1.4 Dynamic pricing

Another way firms can compete or increase revenue is by setting prices dynamically over time. In the traditional structure, the physical changes of price stickers are difficult. You can still find the situation today in stores where the price tag does not match the actual prices you pay. However, e-commerce makes dynamic pricing much easier. You may experience this and have tried to benefit from this. In fact, price change is so prevalent, it spawns a new market for Apps that search prices frequently and alert their potential buyers. It has also led to research about online learning.

2. AUCTIONS

An auction is a process of selling or buying products or services through a bidding process. You may have seen the auction of livestock, antiques, artwork, or used cars (Manheim) live or in movies. A fast-talking auctioneer calls out a starting price. Once there is a bid, the auctioneer acknowledges the bid, and waits for a higher bid, and repeats it until there is no higher bid, and then calls "going once, going twice … sold to…" at the highest bid. Sometimes, the auctioneer may call the deal off if the highest bid is below the reserve price, or a preset minimum. This is called an English auction, which we may be most familiar with. You may have used online auctions on eBay, Priceline, and Lending Tree. Online advertisements on Google are auctioned automatically by algorithms behind the scene. In auctions, the buyer or the seller sets their private values and then try to get the best deal through bidding.

2.1 Types of auctions

There are many types of auctions. They can be for selling or buying products or services. Antique or eBay auctions are to sell antiques. Google uses auctions to sell space for potential advertisers. In the supply chain, firms can also set up auctions to buy supplies or services. When an auctioneer wants to buy something, it is called a reverse auction.

The auction can also be open or sealed. An antique auction is open auction in which everyone can see the bidding process openly, such as in an English auction. The closed auction, or sealed auction, such as a silent auction, is not open until to the time when the winner is announced. Each bidder has a single chance to bid. Bidders cannot see each other's bid.

In a supply chain, a firm may want to sell facilities, products, or services. For example, a manufacturer may need to sell its product to a wholesaler. A manufacturer may also need to buy the transportation service for its inbound supplies. The firm would prepare a call for proposals (CFP). For example, Texas Instruments may need a distributor to sell its products in a region. The CFP would spell out the details of volume, time, paperwork requirements, and so forth. The potential distributors can bid to get the business.

Texas Instruments may also need someone to supply its display. The CFP would spell out the specifics pertaining to their design, the supply frequency, quantity, lead time, stockout penalties, and soon. The potential display suppliers can bid to provide the supply within the deadline. This is considered a reverse auction.

In both cases above, Texas Instruments is considered the auctioneer. The distributors or suppliers are bidders. The CFP must specify how the firm determines the winner, such as the highest price

or lowest charge. Potential buyers or sellers will send in their bids or charge. The firm will determine which bidder based on rules that everyone knows.

In supply chain, the bidders cannot see each other's proposals. Therefore, this is a closed or sealed bid auction. Each bidder may not value the proposal at same value. For example, a city is auctioning off a piece of commercial land. The potential buyers can send in their offering bid. The value of the property for each bidder varies depending on the type of firms and their potential use of the land. This is called private value.

Government organizations, such as NASA, the Department of Defense, and the State of Georgia, are normally required to accept the lowest bid when they need products or services. If the lowest-bid supplier is not selected, the agency must show strong evidence why the lowest bidder was not selected.

2.2 Private value

Both the auctioneer and the bidder have some predetermined value of the product or service. The city that wants to sell the parcel of land appraises the parcel and determines its value to the city before holding the auction. The bidders would do their independent research to determine their values. However, the same parcel may have different values to different parties. We call this private value.

As a client in a supply chain, a firm can determine the private value based on the benefits of winning the auction, such as additional profit it can generate. It should also consider the synergy of the new acquisition with the existing business. For example, Texas Instruments would investigate how much leverage the new screen supplier can provide in relationship with other needs.

The private values among auctioneer and potential bidders normally fall into a range. For example, the private values among all involved may fall between one and two million.

3. SEALED-BID, PRIVATE VALUE AUCTION IN CLASS

A sealed-bid, private value auction is the most relevant between supply-chain partners. Before we go to the theoretical analysis, we can experience the sealed bid private value auctions. This can be done through some platform, such as MobLab, or in class using clicker entries.

3.1 In-class games

This can be done using some classroom response system such as TurningPoint, or in a platform such as MobLab. An item with its relevant information will be shown to the students. Each student can determine his or her private value for the item. This is similar to industrial auction of a parcel of land, a distribution center, or transportation services. Once started, each student can bid some value based on his or her private value and strategy. Students cannot see each other's bid values. When bidding ends, the student with the highest bid wins the item. He or she will pay the bid value and get the item. If tied, breaking the tie will follow predetermined tie-breaking rules.

3.1.1 Cross pen in-class Auction record

The first game uses the first-price rule, meaning the bidder with the highest bid wins the auction and pays the bidding price. The results are shown in Figures 5.1 and 5.2. The initial bid was just

for the practice run. You can see that the average bidding price is about 10% higher than the private value. Nice, this is not a big deal! Imagine if you are purchasing a million-dollar parcel of land!

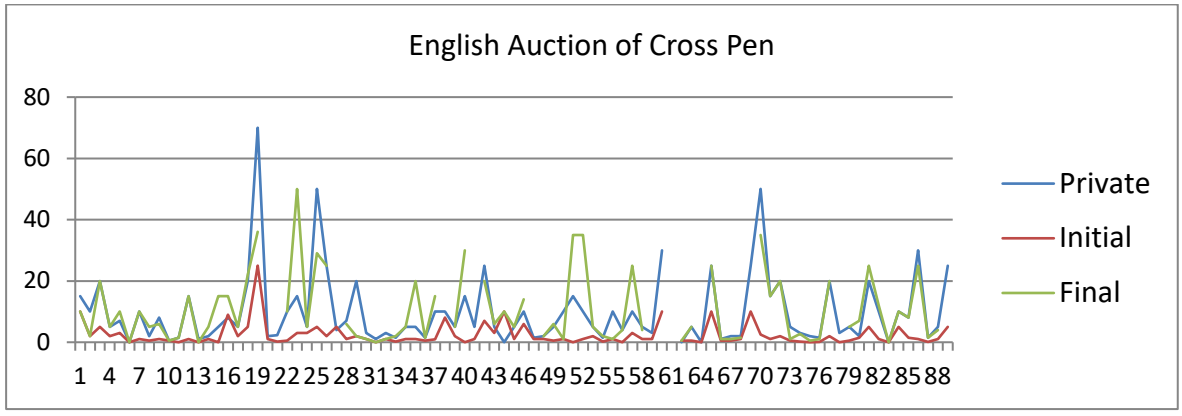


Figure 5.1. Individual private value, initial and final bids on a Cross pen.

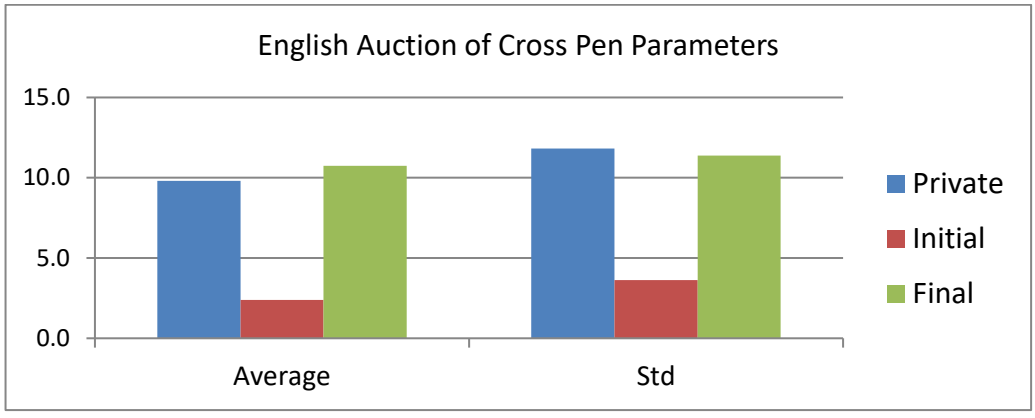


Figure 5.2. Summary data on private value, initial and final bids of Cross pen.

3.1.2 Vickery or second price auction

William Vickery is a Nobel laureate in Economics. One of his contributions is the second-bid auction. In a second-bid auction, the bidder with the highest bid wins the bid. However, he only has to pay the second-highest bid to claim the prize.

Vickery proved that in a second-bid auction, the bidders are incentivized to bid their true values, instead of bidding lower than private values.

3.1.3 Summary of in-class auctions

Private value varies because of information, personal interest, synergy with what the bidder already has, or attachment. In first-price auction, the bidders would bid below their private values because if they bid at their private values, the utility is zero, the same as losing or not participating. If they bid higher than their private value, even when they win, they lose in utilities. People are susceptible to bidding above the private value in an open auction. In the excitement of the bidding process, and the influence of other bidders' bids, one may let the environment push their bids above their preset private value. This is called the winner's curse. The question is how much below the private value should one bid?

In the second-price auction, bidding at above the private value may still leads to positive utilities if the second highest bid is below the winner's private value. What is the best strategy, then?

4. MODELING AND ANALYSIS OF PRIVATE VALUE SEALED-BID AUCTIONS

Buyer and seller both have private value: how much it is worth to you. Theoretically, when the buyer's private value is higher than seller's a transaction can be reached. In a supply chain, the value of products and services depends on many other factors, such as the financial situation and the relationships of other products and services each player has.

4.1 Modeling of first price auction

Assumptions

1. There is one auctioneer.
2. There is one item to be auctioned.
3. There are n bidders, $n > 1$.
4. The bidder with the highest bid wins the auction.
5. The winner pays its bid value.
6. All bidders are risk neutral, to optimize their utilities.
7. No one knows the other's private value.
8. There exists a common knowledge that the private values of all participants fall into a range.
9. The range of the private values, v's, is identically distributed and follows a continuous uniform distribution $U(0,1)$. The density and distribution are $f(v) = 1$ and $F(v) = v$ between (0, 1), and zero otherwise.

Many of these assumptions can be relaxed. For example, instead of selling, in a reverse auction to buy there can be multiple items; some may have some knowledge on someone's private value, generic uniform distribution $U(A,B)$, nonuniform distribution, and the winner pays the second highest bid value.

Definitions

v – private value of the bidder

b – bidder's bid

n – number of bidders, $i = 1, 2, \ldots n$

$P(b)$ – probability of winning with bid b

For bidder i, with a private value of v_i and bid b_i, there is no reason to bid above the private value because the utility is negative, worse than losing a bid. For a rational bidder, we have

$$b_i < v_i$$

The bidder i's utility depends on if the bidder wins or not (the tie is not considered in this simple case):

$$\pi(b_i) = \begin{cases} (v_i - b_i)P(b_i), & \text{If bid } i \text{ is the highest} \\ 0\big(1 - P(b_i)\big), & \text{if the bid } i \text{ not the highest} \end{cases}$$

Where $P(b_i)$ is the probability of $b_i > b_j \; \forall i \neq j$.

Figure 5.3 shows the relationships between the winning margin, probability of winning, and the bidder's utility. The winning margin is $v - b$, a declining linear function of b. The probability of winning is an increasing function. If there are two bidders, this is a linear function. For three players, this becomes quadratic. As the bid increases, the probability of winning increases. The utility is the dashed line, or the product of winning margin and probability of winning increases, at decreasing rates. At some point, the utility starts to decrease. When the bid is the same as the bidder's private value, the utility is zero, same as losing. As the bid increases more, the utility becomes negative due to negative margin.

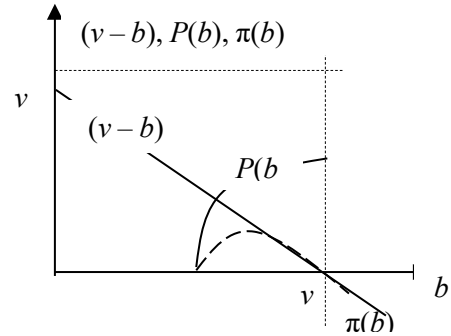


Figure 5.3. The relationships between the winning margin, the probability of winning, and utility.

Therefore, a good bid should have a positive margin below the bidder's private value. How much margin is the best response? What does the best response margin depend on?

4.2 Bidder's best responses and utilities when there are two bidders

In games, one's utility depends on other's actions. In auction games, we need to add conditional probabilities, such as $U(0,1)$ to find the solutions. Since the solutions are based on some

distributions as conditions, the results in auction are referred to as a Bayesian Nash equilibrium (BNE). This is similar to the situation in a Cournot game when a player does not know the opponent's marginal cost but has to assume some distribution of the marginal cost.

4.2.1 Best bidder response

We have established that a bidder should bid $b < v$. If both bidders are rational, they would both bid at a fraction c of their private values, or $b_1 = cv_1, b_2 = cv_2$. Based on the assumption that the private values follow a uniform distribution $U(0,1)$, the probability that bidder 2's bid is lower than bidder 1's bid value is $P(b_2 < b_1) = P(cv_2 < b_1) = P\left(v_2 < \frac{b_1}{c}\right)$. We can now use the uniform probability $U(0,1)$ as condition to find the best response bid.

$$P\left(v_2 < \frac{b_1}{c}\right) = \int_0^{\frac{b_1}{c}} 1 * dv_2 = \frac{b_1}{c}$$

The utility for bidder 1 if bidder 1 wins

$$\pi_1(b_1) = (v_1 - b_1)P\left(v_2 < \frac{b_1}{c}\right) = (v_1 - b_1)\frac{b_1}{c}$$

Applying first-order condition with respect to b_1, we have

$$b_1 = \frac{v_1}{2}$$

Since the two bidders are symmetric, this best response also applies to bidder 2, or

$$b_2 = \frac{v_2}{2}$$

This means for two bidders, the fraction c = 0.5, or both bidders should bid half of their private values.

4.2.2 The expected utility winner and expected utilities for both

The utility for the winner is

$$\pi(b_{win}) = \left(v_{win} - \frac{v_{win}}{2}\right)\frac{\frac{v_{win}}{2}}{0.5} = \frac{{v_{win}}^2}{2}$$

This is a quadratic increasing function of the bidder's private values.

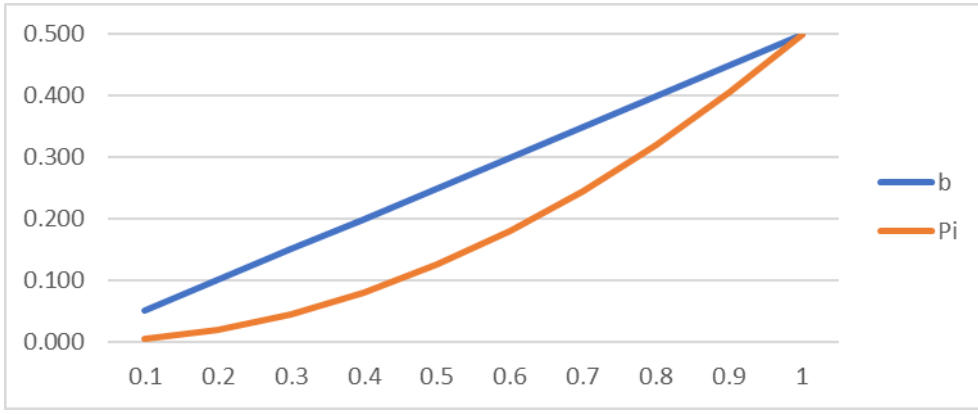


Figure 5.4. The bid value and expected utilities for the winner.

For a known private value, this depends on the actual value. The expected utility for both bidders is only a function of number of bidders, or

$$E[\pi, 2] = \int_0^1 \frac{v^2}{2} dv = \frac{1}{6}$$

Which does not depend on the private values.

4.2.3 The revenue for the auctioneer and utility for two bidders

The probability of the auctioneer's expected revenue can be found in two possibilities. If $b_1 > b_2$, bidder 1 wins with a winning bid of $b_1 = 0.5v_1$. The probability for this to happen is the probability of $v = v_1$, or $f(v_1)$, times the cumulative probability of $v_1 > v_2$, $F(v_1)$. We have

$$f(v_1)F(v_1) = 1 * F(v_1) = 1 * v_1 = v_1$$

Similarly, if $b_2 > b_1$, the probability is v_2. Since they are from the identical distribution, we can combine and drop the subscript and sum them up. $f(v)F(v) = 2v$. Then, the private value of bidder with winning bid is

$$E[v_{win}] = \int_0^1 v * 2v dv = \frac{2}{3}$$

You can also find that the private value of the losing bidder is $\frac{1}{3}$. The average of private values of winner and loser is 0.5, the expected value of $U(0,1)$.
Since the rational bidder bids half of the private value, the auctioneer's revenue when there are two bidders is

$$R(b) = \frac{1}{2}\frac{2}{3} = \frac{1}{3}$$

Since the expected private value of the winner $v_{win} = \frac{2}{3}$, we can also find the expected utility for the winner before we know the private value, or

$$\pi(b_{win}) = \frac{{v_{win}}^2}{2} = \frac{\left({}^2/_3\right)^2}{2} = \frac{2}{9}$$

You can also find

$$\pi(b_{lose}) = \frac{{v_{lose}}^2}{2} = \frac{\left({}^1/_3\right)^2}{2} = \frac{1}{18}$$

The expected utility of both is

$$E[\pi, 2] = \frac{\pi(b_{win}) + \pi(b_{lose})}{2} = \frac{1}{6}$$

This is the same as we derived directly earlier.

4.3 Best responses and utilities when there are *n* bidders

When there are n bidders, we have established that a bidder i should bid $b_i < v_i$. If all bidders bid at a fraction c of their private values, or $b_j = cv_j | \ 0 < j < 1, \forall j \neq i$. Based on the assumption the private values follow a uniform distribution $U(0,1)$, the probability for one bidder j's value less than bidder i's bid is

$$P(b_j < b_i) = P(cv_j < b_i) = P\left(v_j < \frac{b_i}{c}\right)$$

The probability that i^{th} bidder wins the auction is

$$P(b_j < b_i) | \forall i \neq j = \left[P\left(v_j < \frac{b_i}{c}\right)\right]^{n-1} = \left(\frac{b}{c}\right)^{n-1}$$

The utility for any bidder is

$$\pi(b) = (v - b)\left(\frac{b}{c}\right)^{n-1}$$

Applying the first-order condition, we can find the best response for a bidder, or BNE bid value is

$$b_i = \frac{n-1}{n} v_i$$

Figure 5.5 shows the fractions of private value of the best response bid at different number of bidders. The utility for the winning bidder is

$$\pi(b_{win}) = (v - b)\left(\frac{b}{c}\right)^{n-1} = \left(v - \frac{n-1}{n} v\right) v^{n-1} = \frac{{v_{win}}^n}{n}$$

The expected utility among all bidders is

$$E[\pi, n] = \int_0^1 \frac{v^n}{n} dv = \frac{1}{(n+1)n}$$

This function decreases quickly with number of bidders, because most bidders will not win and have zero utilities.

n	2	3	...	5		10	...	50
c	0.5	0.667		0.8		0.9		0.98
$\pi(b)$	$0.5v^2$	$0.33v^3$		$0.2v^5$		$0.1v^{10}$		$0.02v^{50}$
$E[\pi, n]$	0.167	0.083		0.033		0.009		0.0004

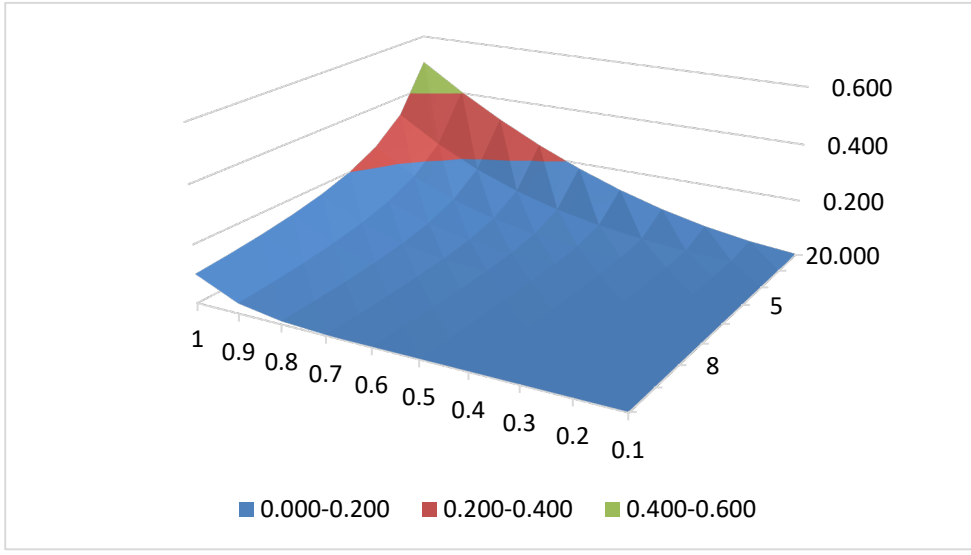


Figure 5.5. The expected utilities for *n* bidders.

The best responses as a fraction of the private value c increase with number of bidders. This makes sense, because the more bidders there are, the more likely someone may have higher private values. The utilities for a bidder depend on $(1 - c)$ and the private value raised to the power of number of bidders. Since the private value is less than 1, this term decreases rapidly with the number of bidders.

When there are n bidders, the probability for a bidder with the highest bid is v^{n-1}. There are n different ways. Therefore, up $f(v)F(v) = nv^{n-1}$. Each bidder bids at $\frac{n-1}{n}v$. Therefore, the expected revenue among all bidders for the auctioneer is

$$R(n) = \int_0^1 \left(\frac{n-1}{n}v\right)(nv^{n-1})dv = \frac{n-1}{n+1}[v^{n+1}]_0^1 = \frac{n-1}{n+1}$$

The auctioneer's expected revenues as the number of bidders are shown in the table below and in Figure 5.6.

n	2	3	…		5	…	10	…	50
Revenue	0.333	0.5			0.667		0.8182		0.96

Graphically, the expected revenue

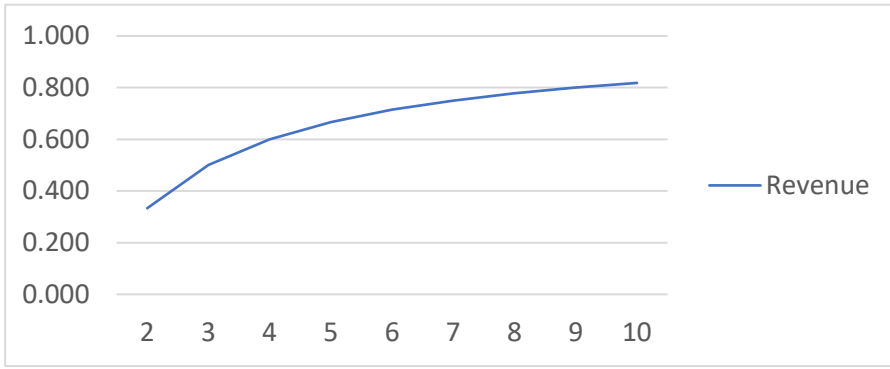


Figure 5.6. The expected revenue for the auctioneer versus number of bidders.

This is an increasing and concave function of number of bidders. As $n \to \infty$, $\Pi(b) \to 1$. The utility for the auctioneer is

$$\Pi(n) = \frac{n-1}{n+1} - v_{auctioneer}$$

The auctioneer's utilities grow at higher rate than the revenue growth because auctioneer's private value is a constant. The auctioneer should try to find more bidders!

4.4 Summary basic model with *U*(0, 1) uniform distribution

In a sealed bid, private value auctions when the private values of all n bidders follow uniform distribution $U(0,1)$, the BNE response for bidder i is to bid the $b_i = \frac{n-1}{n} v_i$. As n increases, the fraction of the private value decreases, and approaches zero asymptotically. The expected utility for the winning bidder is $\pi(b_{winner}) = \frac{{v_{winner}}^n}{n}$, which decreases with the number of bidders rapidly. The expected utility among all bidders is $E[\pi] = \frac{1}{(n+1)n}$, which decreases with the number of bidders even more rapidly because more and more bidders do not win.

For the auctioneer, the revenue with n bidders is $R(n) = \frac{n-1}{n+1}$, which increases rapidly and then slows and asymptotically approaches 1, the upper limit of the uniform distribution.

4.5 Extension to *U*(*A*, *B*) uniform distribution

In practice, the private value range may be from some minimum to some maximum. For example, a distribution center may be worth between 1 and 2 million dollars. This is a simple transformation

to $U(A,B)$. We use the capital A and B to distinguish the bid value b from the lower limit of the uniform distribution.

If the lowest possible private value is A, there is no need to bid below A. The best response should be the fraction of private value minus the lower limit plus the lower limit, or

$$b_i = A + \frac{n-1}{n}(v_i - A)$$

Where, $A \le v_i \le B$. The expected utility for the winning bidder is

$$\pi(b_{win}) = \left((v_{win} - A) - \frac{n-1}{n}(v_{win} - A)\right)\left(\frac{v_{win} - A}{B - A}\right)^{n-1} = \frac{(v_{win} - A)^n}{n(B-A)^{n-1}}$$

The expected utility among all bidders is

$$E[\pi, n] = \frac{(B-A)}{(n+1)n}$$

The expected revenue for the auctioneer is

$$R(n) = A + \frac{n-1}{n+1}(B - A)$$

In summary, with the uniform distribution $U(A,B)$, we must shift the lower limit to A, and adjust the range to $B - A$.

Example

A firm is to sell a distribution center with all its facilities. The firm believes that the distribution center is worth 3.1 million. There are two potential buyers: Ann and Bob. The private values for the firm and the two potential buyers are between 3 and 4 million. Ann studied the building and facility and discussed them with her colleagues. They concluded that the facility is worth 3.8. Bod did the same with an evaluation of 3.6 million. Please find the following:

1. The bidders' BNE bids.
2. The bidders' expected and actual utilities.
3. The firm's expected and actual revenue and utility.
4. If $n = 9$, the expected utilities for all the bidders and the winner.
5. If $n = 9$, the expected revenue and utility for the firm.
6. Observations from this example.

Assumptions

All the modeling assumptions hold, except the last assumption. We assume the private values are iid and uniformly distributed $U(3,4)$.

Solutions

1. The bidder's BNE bids

The bidder's BNE bids is half of the private value, or

$$b_{Ann} = b_{Ann} = A + \frac{n-1}{n}(v_{Ann} - A) = 3 + \frac{2-1}{2}(3.8 - 3) = 3.4$$

$$b_{Bob} = 3 + \frac{2-1}{2}(3.6 - 3) = 3.3$$

2. The bidders' expected and actual utilities:

The expected utilities before bidding:

$$\pi_{Ann}(3.4) = \frac{(v_{Ann}-A)^n}{n} = \frac{(3.8-3)^2}{2} = 0.32$$

$$\pi_{Bob}(3.3) = \frac{(v_{Bob}-A)^n}{n} = \frac{(3.6-3)^2}{2} = 0.18$$

The actual utilities after bidding

$$\pi_{AnnActual}(3.4) = v_{Ann} - b_{Ann} = 3.8 - 3.4 = 0.4\ mil$$

Bob's utility is 0.

3. The firm's expected and actual revenue and utility

The expected revenue before the bidding using the BNE formula

$$R(2) = A + \frac{n-1}{n+1}(B - A) = 3 + \frac{2-1}{2+1}(4 - 3) = 3.33\ mil$$

The expected utility for the firm is then the expected revenue minus its own utility, or

$$\Pi(2) = R(2) - v_{Firm} = 3.33\ - 3.1 = 0.23$$

The actual utility for the firm is the difference between Ann's bid and the firm's private value, or

$$R_{Actual}(3.4) = b_{Ann} = 3.4$$

$$\Pi_{Actual}(3.4) = b_{Ann} - v_{Firm} = 3.4 - 3.1 = 0.3$$

4. If $n = 9$, the expected utility for all 9 bidders before the auction

$$E[\pi, 9] = \frac{(B-A)}{(n+1)n} = \frac{4-3}{(9+1)*9} = \frac{1}{90} = 0.0111 \ll E[\pi, 2] = \frac{1}{6}$$

The expected utility for the winner depends on the winner's private value

$$\pi_i(v_i) = \frac{(v_i-3)^9}{9} = 0.111(v_i - 3)^9$$

We can find this in a different way from the expected utility among all 9 bidders. For 9 bidders, there are 8 who got zero utility. Therefore, we have

$$\pi(b_{win}) = nE[\pi, 9] = 9 * 0.111 = 0.1$$

5. If $n = 9$, the expected revenue and utility for all the firm are

$$R(9) = A + \frac{n-1}{n+1}(B - A) = 3 + \frac{9-1}{9+1}(4 - 3) = 3.8$$

$$\Pi(9) = R(9) - v_{Firm} = 3.8 - 3.1 = 0.7$$

6. Observations from this example.
    a. The actual utilities depend on actual bids (Bayesian): uncertainty, uniform, or other
    b. Provide estimates and strategies
    c. Losing bid does not impact
    d. The bidder's expected utility decreases fast with number of bidders and asymptotically to zero
    e. The auctioneer's expected revenue increases with number of bidders, asymptotically toward upper bound.

4.6 Vickery second-price sealed-bid private value auction

Vickery proposed the second-price sealed-bid auctions. All the assumptions remain the same. However, the mechanism of pay is changed. The bidder with the highest bid wins the auction. However, the winner will only pay the second-highest price. He proved that in the second-price sealed-bid auction when the bidders' private values follow uniform distribution, the best response is to bid the private value, or

$$b_i = v_i$$

This strategy maximizes the chance of winning for each bidder and ensures some margin since the second-highest bid is lower than the winner's private value. This result is interesting.

Recall, in the first price auction with two bidders, the expected winner's private value was 2/3.

$$E[v_{win}] = \int_0^1 v * 2v dv = \frac{2}{3}$$

Ironically, the bidder's utility from the best response and the auctioneer's revenue remains the same. If you think carefully, this makes sense based on the uniform distribution assumptions. Based on the uniform distribution, the expected private values are uniformly distributed between the low and high. The expected second bid is $v_{2nd} = \frac{n-1}{n} v_{1st}$. Therefore, the auctioneer's revenue is the same for the first price and second price auctions. In practice, the bidder's private values may not follow uniform distribution and therefore, the auctioneer's revenue in second price auction can be higher or lower than in the first price auction. The game designers study the situation and select the winning mechanism.

Example

Consider the same example of a firm auctioning off a parcel of land. Ann and Bob are the bidders. The rule is changed to second price. The private values are estimated to follow $U(3,4)$ by all parties. Other details $v_{Firm} = 3.1$, $v_{Ann} = 3.8$, $v_{Bob} = 3.6$. Please find

1. The BNE bid for each.
2. Their expected utilities of the bidders and the revenue for the firm before auction.
3. The actual utilities for both bidders and the revenue for the firm.
4. What is the firm's expected and actual revenue and utility if Ann and Bob bid their BNE?
5. What is the expected revenue, and provide for the firm if there are nine bidders?

Solutions

1. The BNE bid for each

Ann and Bob would their true private values, or

$b_{Ann} = 3.8$, $b_{Bob} = 3.6$

2. The expected utilities if Ann and Bob bid their BNE before bidding

The expected for Ann remains the same at 0.32.

The expected utility for Bob remains the same at 0.18.

The firm's revenue before auction remains the same at 3.333

3. The actual utilities for all

The actual utility for Ann is that Ann will pay Bob's bid at 3.6. Therefore, Ann's utility is 3.8 – 3.6 = 0.2, 0.12 million below the expected. This is due to a higher-than-expected private value for Bob.

Bob lost. His utility stays at 0.

The revenue for the firm is 3.6.

4. What are the firm's expected revenue and utilities based on the uniform distribution?

The firm's revenue expected and utilities remain the same at 3.3.

4.7 Summary of sealed-bid private-value auction

The theoretical analysis showed several interesting insights into the sealed-bid private-value auctions.

Under the first price rule, the best strategy is to bid some fractions of the private value. The fraction should increase with the number of bidders. The bidder's utilities decrease fast with the increasing number of bidders.

The auctioneer's revenue increases with the number of bidders.

Under the second-price rule, the best strategy is to bid at the private value. However, the expected pay, and the expected revenue for the auctioneer are the same if the private values follow a uniform distribution. The auctioneer can also benefit more from a larger number of bidders.

These results apply to reverse auctions with modifications. The bidders compete to sell products or services. In that case, they would bid higher than their private values.

In supply-chain practice, there is much more at stake. The bidders study to find their private values, considering many factors. For example, when trucking companies bid to serve transportation lanes, each trucking company would consider their current capacity, their resources, the compatibility of the new lanes with their existing lanes, and so on. Competitors can figure out some of the private information through study and use the result to formulate their bidding reality. There is much more gamesmanship. Therefore, the exact results from BNE may not be accurate, as you have experienced in your own auction games. However, the BNE provides great insights for formulating the strategy.

5. WINNER'S CURSE

In open auctions, bidders can let emotion and influence from others lead a bid above its private values. Bidder can also over-bid in common value auctions. In certain items such as an oil field, timber, and broadcast rights, they have the same value no matter who wins the bid, except no one knows for sure what the value is before engaging in the economic activities. Since the winning rule is to select the highest bid, which is biased even when the bids are from an iid, these are referred to as the Winner's Curse. In analysis, this is called an Order Statistic. In essence,

If $X_1, \ldots X_i \ldots X_n$ are n independent and identically distributed random variables. The n^{th} Order Statistic is

$$X_{(n)} = max\{X_1, \ldots X_i \ldots X_n\}$$

If X_i follows $N(\mu, \sigma)$ the bias for n bidders, WC_n is

$$E\left[X_{(n)}\right] = E[X_i] + WC_n\sigma$$

n	1	2	3	5	10	20	50
WC_n	0	.56	.85	1.16	1.54	1.87	2.25

For uniform distribution, the winner's curse follows a beta distribution

$$U_{(n)} = B(n, 1)$$

5.1 Interesting World Cup coverage right bidding in Singapore, 2010

Three companies in Singapore bid for different types of World Cup (Star Hub, SingTel, and M1) broadcast rights. Each overbid by a lot. They considered the price demand curve without knowing the parameters of the curves. They faced a dilemma: High price with low volume or low price with high volume. Since they do not have historical data to use, they have to do it by trial and error. They charged $70 for home use, one small TV in a bar is $3,000, and a large one is $5,000. The results were low viewership and a lot of complaints. They lowered the prices later to reduce the complaints, but each firm lost a lot. Winner's Curse.

6. PRICE DISCRIMINATION

Price discrimination or price differentiation refers to charging different prices for basically the same products or services. We will use the term "discrimination" to differentiate the contents of a Bertrand competition in price differentiation. There are two major flavors of price discrimination:

1. Direct price discrimination. It refers to charging different customers or clients different prices based on observable characteristics of the customer. For example, the prices can be set differently for
   - Students
   - Senior citizens
   - Members such as gym members, IISE members
   - Corporate buyers
   - …
2. Indirect price differentiation is available to all regardless of identity. It is up to the customers or client to decide which price options to take. Examples can be
   - Quantity discount
   - Timed promotions, you have seen in Barilla and Campbell's cases.
   - Coupons
   - Bundling

- Phone plans
- Airline tickets or hotel rooms purchased full fare, non-refundable, or time of purchase,
- …

You have experienced MobLab price discrimination. MobLab publishes the charges on their website (Figure 5.7). In addition, MobLab offers a student discount for partner universities. The price for the students in the partner universities is $5 per student per course. The university also pays a similar share based on the expected enrollment.



Figure 5.7. MobLab prices (https://moblab.com/edu/pricing, accessed on September 6, 2021).

Are these three prices direct or indirect discriminations? How about the $5 price for the students in partner universities?

One extreme price discrimination is used in hospitals. Hospitals normally use a complex and secretive charge schedule commonly refers to as a "charge master." The prices of the same products or services vary widely among patients with or without insurance, with which insurance, with cash, or in other ways. They often charge lower prices for Medicare or Medicaid patients. You can find a lot of articles attempting to decode how they charge. However, it has been difficult. Part of the difficulty is that the coding in medical services is so complex, there are master's degrees offered just to use the coding to effectively increase revenues. There are crowd-sourced services that can provide some help for the patients to lower their costs.

There is also price discrimination among different countries. An extreme example is medical drugs. Many countries negotiate with the suppliers as a country to enjoy stronger negotiation power. You can search to find the huge differences in drug prices between countries. This is also true for technical textbooks. There is developing-country discount. The cost in drugs in United States is very high. The pharmaceutical firms say the reason is that the cost of research and development is high, and mostly recovered in the U.S. market. You may want to put your critical mind on to think about this.

Price discrimination falls in the broad topics of revenue management. We will explore one example in the direct price discrimination of books.

6.1 Direct price discrimination with books

6.1.1 No price discrimination

First, we consider when there is no price discrimination.

Example

The maximum quantity of a book price in a bookstore is estimated to be $80. For every dollar increase, the demand is estimated to drop by one book. The cost of the textbook is $20. Please

1. Find and sketch the demand curve.
2. Find the price that maximize the profit.
3. Find the expected supply quantity and profit.

Solutions

Normal assumptions in the competition.

1. The mathematical demand curve as a function of price is below. The curve is shown in Figure 5.8.

$$q = 80 - p$$

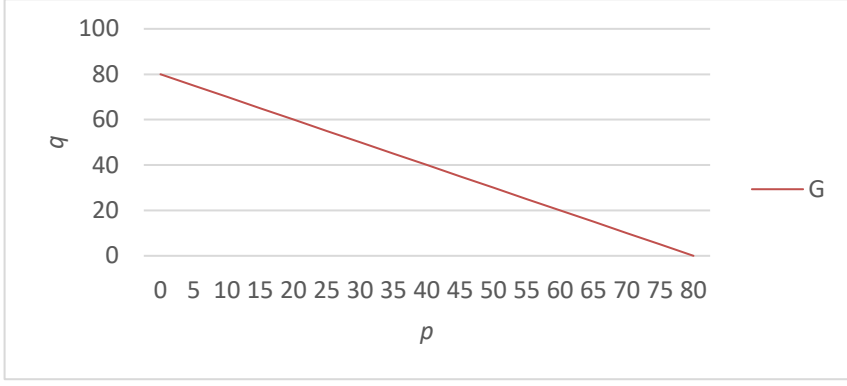


Figure 5.8. The demand function for the book.

2. Find the price that maximizes the profit

Assumptions

The same as we have used in the competition. There are no competitors (very strong).

Solutions

We can write the profit function as

$$\pi(p) = (p - c)q(p) = (p - 20)(80 - p)$$

Apply the first order condition, we have

$$-p + 20 + 80 - p = 0$$

$$p^* = \frac{100}{2} = 50$$

3. Find the expected supply quantity and profit.

The quantity at this price using the demand function is

$$q = 80 - 50 = 30$$

The profit is

$$\pi(50) = (50 - 20) * 30 = 900$$

6.1.2 Price discrimination

A professor adopted the book as a textbook. The maximum demand for the students is 180. The students are more sensitive to the price. For each dollar increase, the demand is expected to drop by three books. Let us repeat the example with the addition of the student market. The demand curve for the students is

$$q_S(p) = 180 - 3p$$

Both the general market and student market demand curves are shown in Figure 5.9.

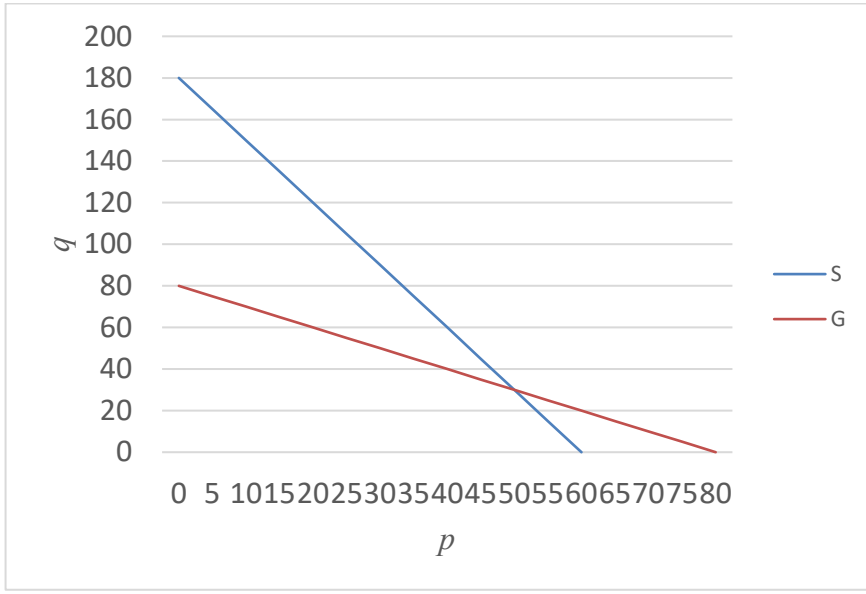


Figure 5.9. The demand curves for both general public and students.

We assume the additional demand will not change the production cost. In practice, this may not be true. The economies of scale can have impact. The profit function for the student market is then

$$\pi(p) = (p - c)q(p) = (p - 20)(180 - 3p)$$

Applying the first-order condition, we have

$$-3p + 3 * 20 + 180 - 3p = 0$$

$$p_S^* = \frac{240}{6} = 40$$

The expected sales in the student market are

$$q_S(40) = 180 - 3 * 40 = 60$$

The profit for the student market is

$$\pi_S(40) = (40 - 20) * 60 = 1200$$

If we assume there are truly two separate markets—no student bought the textbook and no non-student would pay student prices—the total profit will be the sum from both markets, with different prices and individual supply quantities, or \$1,200 + \$900 = \$2,100. This direct price discrimination can be implemented by checking the student IDs.

However, it is possible for cannibalization. Some student may buy the same book at \$50 without the discount, or forgot their ID at the time of purchase, and need the book urgently without worrying too much about \$10 extra cost. Therefore, the estimated profit may be higher than reality.

Software companies often provide a student version. You may have used software in simulation, CAD, and so on. Once the students learned a type of software, they may want to use the same software once they go to work.

6.2 Indirect price discrimination

In Chapter 7, Economies of Scale, we will discuss one form of indirect price discrimination: quantity discount.

7. DYNAMIC PRICING

Dynamic pricing is to change prices overtime. For airlines and hotels, in addition to their routine Saturday night stay discount, they may also do additional changes dynamically to reflect the market and competition. Amazon varies their prices frequently without prior notification. The internet has provided much better infrastructure to perform dynamic pricing. Firms can create automatic searches on prices of the competitors to set their prices dynamically. The firms can study buyer's behavior through their search history and click patterns. They can use such information to set prices to maximize their revenues. There are courses that focus upon online learning. One of the applications of online learning is to set prices dynamically.

8. PRICING AND HUMAN NEEDS

Auction is an efficient way to set prices to match the firm's needs. Price discrimination can be used to help the needy. The quantity discount can be used to be competitive by passing the savings from economies of scale to their customers. Price discriminations by airlines and hotels make travel more affordable. Travel is one of the best ways to increase the satisfaction of human beings. These are all great.

However, revenue management simply trying to squeeze more out from the clients may not always serve human needs. Quantity discount to encourage over-consumption or waste can increase the environmental footprint. Food over-consumption is also bad for human health, a basic human need!

As supply-chain professionals, should we look beyond the top line and bottom line?

6. LOGISTICS IN SUPPLY CHAIN

Logistics is an important part of the supply chain that moves and stores the goods for efficient flow throughout the supply chain. CSCMP defines logistics management as the part of supply-chain management that plans, implements, and controls the efficient, effective forward and reverse flow and storage of goods, services, and related information between the point of origin and the point of consumption in order to meet customers' requirements. Logistics management activities typically include:

1. Inbound and outbound transportation management, and fleet management.
2. Warehousing, materials handling, and order fulfillment.
3. Logistics network design.

Manufacturers, distributors, and retailers may use logistics services. Logistics service providers provide these services. It is important for supply-chain professionals to have good understanding of logistics.

1. MAJOR COSTS IN LOGISTICS

The flow of goods is similar to the flow of water from sources to us. Reservoirs, gates, pipes, and tanks are used to supply the right amount water at the right locations at the right time. There is flow and storage. With discrete goods, the flows are through transportations and storage through warehouses, distribution centers, store shelves, and so forth. Therefore, the major costs in logistics include transportation, inventory, and administrative costs. From 1981 to 2005 the costs are shown in Figure 6.1. During this time, GDP quadrupled.

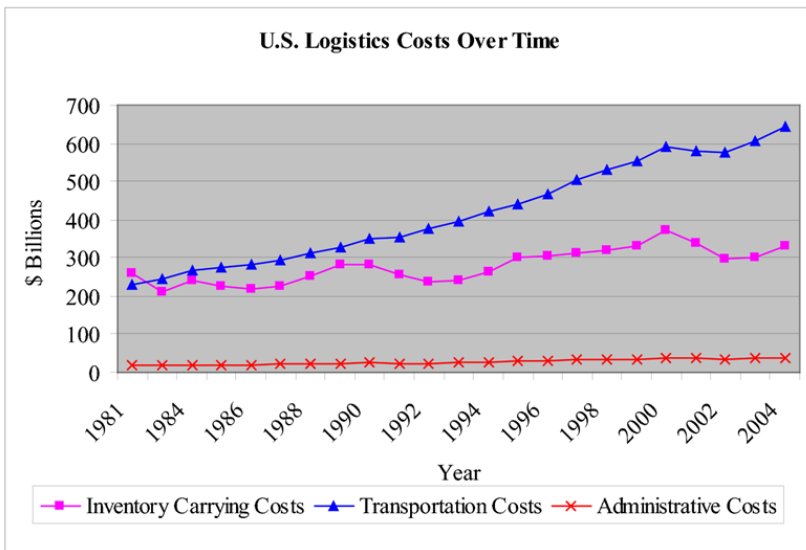


Figure 6.1. Logistics costs, 1981–2005.

During this time, the transportation cost increased most. In 25 years, it tripled, almost in step with GDP growth. This is during a time with significant reduction in manufacturing. Transportation costs are sensitive to fuel cost, driver shortage, and road congestion. Most of the transportation is done by trucks in United States. A truck is a unit load that needs a truck and a driver. This cost grows with inflation. Due to many reasons, such as during the pandemic, the driver shortage often increased the transportation cost.

The inventory cost increases the slowest. It is achieved by the supply-chain professionals and industrial engineers over the years and related to the lean effort.

The administrative cost increase is also slower than the transportation cost but with much better services, thanks to computers and IT. In fact, the administrative cost helped to reduce the transportation cost and inventory cost.

2. TRANSPORTATION

Transportation cost is high and increases with the rest of economy. The cost of the driver, the truck, the fuel, the insurance, and maintenance increase with growth and inflation. The total cost depends heavily on the mode of transportation.

2.1 Transportation modes

Goods can be transported in different modes: over water on ships or barges, on rail carts, on trucks, or through the air. Oil and gas can also be transported in pipelines. The modes have distinctive characteristics in terms of cost, speed, and reachability.

2.2 ISO containers

One of the common unit loads among different modes is containers. The ISO container is an important enabler for the globalization. Containers are standardized unit loads that protect the goods and facilitate fast handling and easy storage. They can be loaded and unloaded quickly on ships, train carts, or trucks with cranes or forklifts. They can be stacked up or retrieved quickly.

Malcomb Mclean developed containers after World War II. Today, ISO defines the width, two heights of 8 or 10 feet, and two lengths of 20 and 40 feet. On the road, train, and ships, most containers are 40 feet long. Note that the weight limit for the 40-foot container is only slightly higher than that of the 20-foor containers, making them more suitable for lighter goods. The standard measure of volume on ship or train is twenty-foot equivalent unit (TEU). The book *The Box* by Mark Kevinson captured the history and how it revolutionized the global trade and sped up development around the world.

In the United States, there are non-ISO standard containers used in intermodal transportation. More on that later.

2.3 Waterborne freight

Transportation on water incurs a low cost. The cost of cross-continent container shipping on large vessels is on the order of one cent for one ton-mile! In addition to ocean shipping, there is also water freight on the Great Lakes and other inland waterways in United States. There are vessels and barges for containers, oil, and bulk commodities. Consumer goods are mostly transported in containers. The largest container ship in 2020 could hold up to 24,000 TEUs! It would take 12,000

trucks, or almost 100 trains to unload one such ship. If you line up the containers back-to-back in single file, it will be almost 100 miles long!

However, ocean shipping is slow. It would take two weeks from Europe to the East Coast, and more than three to cross the Pacific. If the ship must pass through the Panama or Suez Canals, the ships have to wait in the queue to get call time. For international shipping, there are additional delays in customs, security checks, and so on.

In addition, few firms have access at the port, other than in Norway. The goods must be transported to the client via drayage services, which adds to the delays and additional costs in both transport and storage.

Currently, container shipping is highly unbalanced between Asia and North America, and between Asia and Europe. As a result, the cost of shipping can vary a lot depending on the direction of transportation and the seasons.

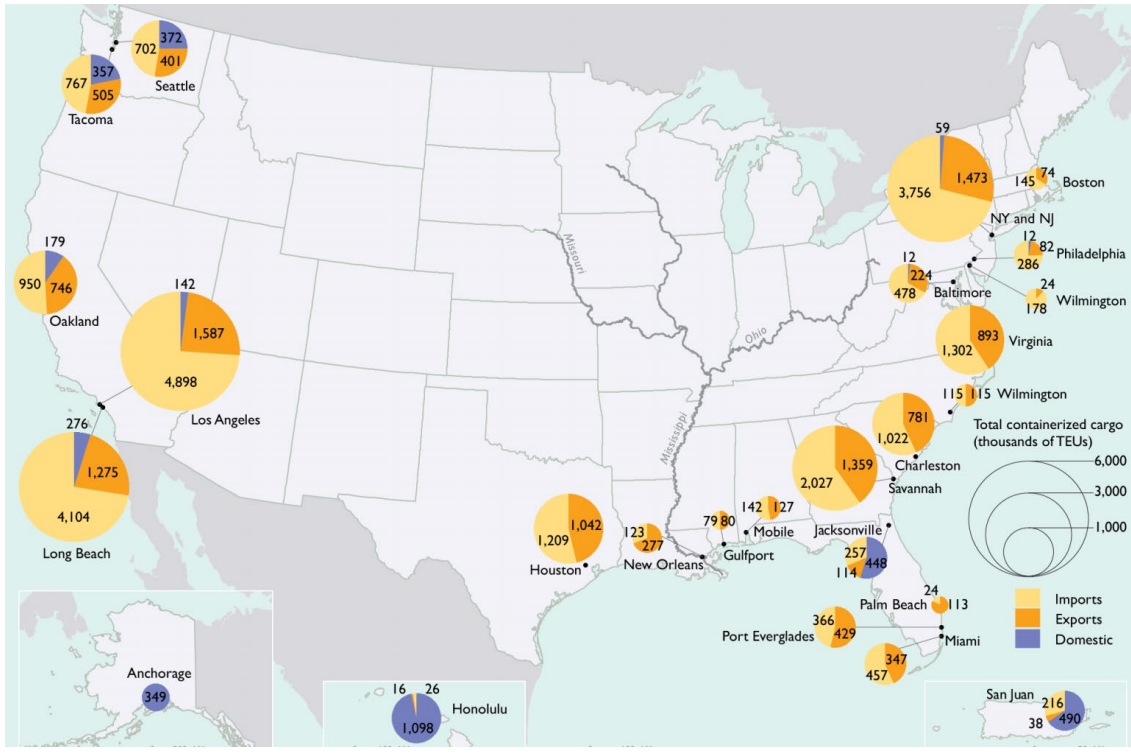


Figure 6.2. The number of TEUs in the top 25 ports by import, export, and domestic use in 2018, by U.S. Bureau of Transportation Statistics, https://rosap.ntl.bts.gov/view/dot/43525, 7/1/2020, PDF file page 7.

2.3 Rail

Rail offers low-cost operations on land. Trains consume less energy than trucks because steel wheels on steel tracks create less friction than rubber on the asphalt. The loads in a freight train with three engineers is equivalent to the loads more than 100 trucks! It is so inexpensive, that in United States, rail companies can compete with trucks successfully in long distances even riding on their own tracks and using their own classification yards and stations built on their own lands. On the other hand, trucks ride on roads built by tax dollars on public lands. The truck license fees paid for by the trucks are not sufficient for the construction of the roads. Many people believe as the economy and population expand, the road capacity will be maxed out. However, there is still capacity on many parts of the rail system.

Rail transport is slow. There are several reasons. First, a train moves slower than a truck in most of the old tracks, because they were not built for high-speed and heavy loads. In addition, a train contains around 100 train cars. Few origin-destination pairs need a trainload of cargo regularly. Therefore, the train may need to be disassembled and reassembled at a hump yard or classification yard. For example, one-third of a train going from Atlanta to Boston may be reassembled at Richmond, Virginia, and then in Washington, D.C. The classification of sections of trains hundreds or even over a thousand feet long takes long time. Finally, very few origins and destinations have rail access. It will require drayage by trucks. The handling and staging will further delay the process. In the United States, there are five main rail operators: BNSF, CSX, Kansas City Southern, Norfolk Southern, and Union Pacific. Since they have different coverage areas, they do not pose direct competition for most part. In fact, there is a lot of collaboration. For example, Norfolk Southern has little coverage on the West Coast. It has to collaborate with BNSF or Union Pacific to serve its customers, and vice versa.

Figure 6.3 shows the freight transportation by tonnage in waterways, rail, and trucks in different parts of the United States (Bureau of Transportation Statistics, 2017). Trucking is very important throughout the country. Rail is prevalent in certain parts of the country, while the domestic waterways are very important along Mississippi River and its tributaries.

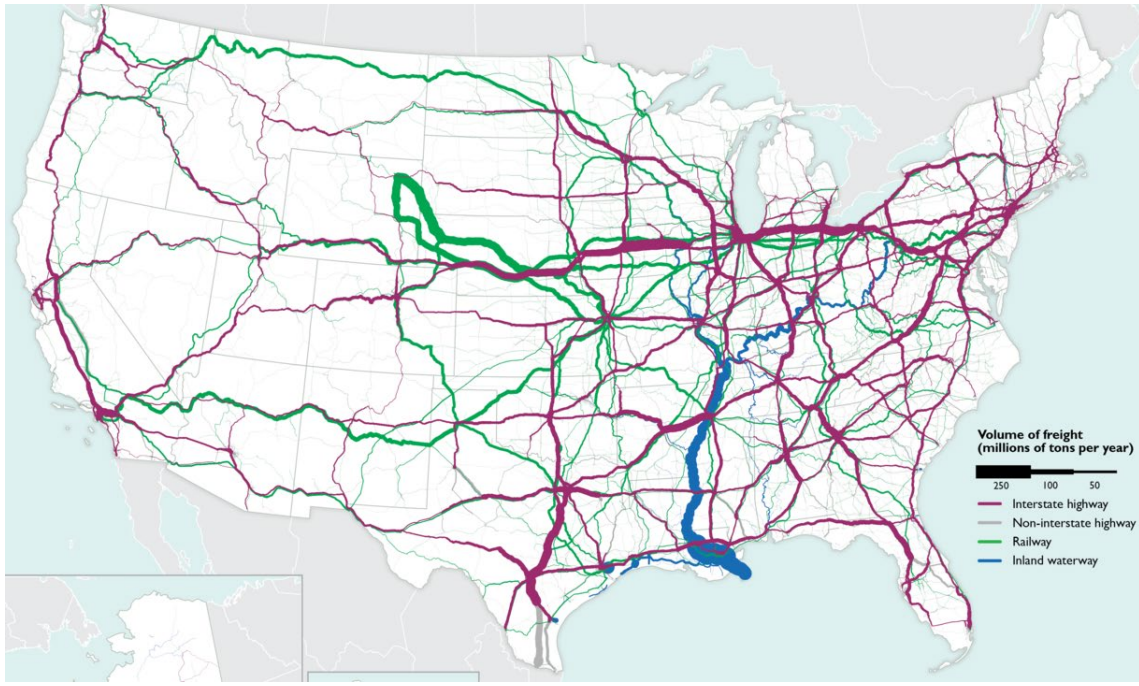


Figure 6.3. Freight tonnage flows by highway, railway, and waterway, 2017, https://www.bts.gov/freight-flows-highway-railway-and-waterway-2017. Accessed July 10, 2020.

2.4 Trucks

In the United States, trucks carry approximately 60% of freight by tonnage and 35% by ton-miles! It is the most important mode in freight transportation. There are many different types of services provided by trucks: truckload (TL), less than truckload (LTL), and package services. There are many different types of equipment used in these services.

2.4.1 Truck load (TL) services

Truckload (TL) truckload services provide point-to-point unit truckload services. The manufacturers need TL services routinely to transport raw materials from suppliers or finished products to distributors or retailers. Over 90% of trucking services are in truckload services. TL provides the lowest cost per ton and per ton-mile and the shortest transportation times. However, often a client does not have enough freight to hire TL services. With the lean and just-in-time drives, the freight service needs become more frequent in smaller quantities.

2.4.2 Less than truckload (LTL) services

The less-than-truckload (LTL) services provide customers with a smaller amount of freight. The weight ranges from 151 pounds (68 kg) up. Once the load starts to approach TL loads, it may be

cheaper and faster to use a TL service. Many LTL services use pallets to load and unload their goods. Some TL services may request their customer to secure their load onto a pallet for easier handling. The challenges for LTL carriers are to maximize the load capacity. Although LTL provides less than truckload services, their trucks are mostly loaded near full. They often adopt hub-and-spoke networks to achieve near truckload with many LTL loads in it. One of the efficiency measures of an LTL service provider is the load efficiency. They can often reach to about 90%. The hub-and-spoke network helped to achiever fuller trucks. However, they need load-change operations en route, which requires more unloading and loading, which leads to delays, potential damage, and getting lost. More on hub-and-spoke networks in the network design chapter after we analyze specific economic forces that drive the success of LTL and related supply-chain network designs.

2.4.3 Package services

Package services are what we consumers often use, such as USPS, UPS, FedEx, or DHL. The packages are easily handled by a single person. Because package services are often personal, each delivery is often unique to a specific address. It is the most expensive service by weight or volume. This is often referred to as last-mile service. Over the years, firms developed many strategies trying to reduce the cost of the last-mile delivery, including network design, high-speed automated sorting systems, and delivery truck design. The most recent development is crowd-sourced services to tap on the idle capacities in the society.

USPS has a Board of Governors and a board-elected Postmaster General, and must report to the Congress but runs like a business. It has the obligation to serve every address in the country, even in remote locations. It must visit most addresses every day and have access to the mailbox. Amazon and Netflix DVD service all utilize USPS effectively due to these characteristics, and to reduce their costs.

2.4.4 Truck equipment

In the Americas, there are roughly three categories of trucks: tractor and trailer, straight trucks, and package trucks. Each has other names. A tractor is the powered vehicle that hauls the trailers. They often come with ten wheels. There is a coupling above the eight rear wheels that can be interfaced with a trailer quickly. A trailer is a load carrier without power. The default trailer in people's mind is a hard-walled box cart with four wheels in the back. The trailer can also be a chassis: a frame with wheels that can be used to carry containers, either ISO standard or other industry standard containers. They can also be flatbed or tankers to carry other types of loads.

The modular tractor-trailer design allows parallel processing. For example, a driver can drop off a trailer at a dock to free up the expensive resources of driver and the tractor for other tasks. The trailer can wait to be unloaded and then loaded again for other loads. The easy separation and engagement offer another advantage: two drivers can drive half a day to meet somewhere in a truck stop to exchange the trailers. This can allow the relaying of the trailer from origin to destination while the drivers can go home every day. The arrangement saves cost on accommodation and makes for happier drivers who can maintain their family life.

The Federal Motor Carrier Safety Administration has rigid rules on hours of service per week. For example, 70 hours per week, with minimum of 34 hours break after, and a 30-minute break

in an 8-hour shift. Many tractors have built in sleeping quarter (sleepers) that the driver can use to sleep at a truck stop, where they can eat and take a shower. Self-driving trucks may change all of this.

In the Americas, the trailers are designed to load and unload from the back via an elevated dock. Almost all factories, distribution centers, and labs have docks. This allows for a fast load and unload for forklifts. For cases, the conveyors can be extended into the trailer for fast load/unload.

The trailer lengths can be 53, 48, or 40 feet long. There is also a 28-foot pup. Two compatible pups can be trained together for flexibility in turning and load-handling. However, the back access requires more yard space and the goods must be first in, last out.

Truck designs in Europe and Japan are different. The engine and cargo space are not easily separable. The engine is normally shorter for compactness. The cargo space has a soft cover. The softcover is less secure than the hard cover as in the Americas. However, the cargo space can be accessed from the side with a forklift. This design allows random access. There is no need for docks. Combined with more a compact design, and shorter turning radius, Europe and Japan can save yard space.

The straight trucks combine the engine and the cargo space for ease of maneuvering. The license is easier to get, and many people can drive. The total length is limited to 40 feet, which gives you up to a 30-foot box. It can hold 18 pallets. There are many different sizes and designs for different purposes.

The package delivery trucks are designed to optimize for the convenience of pick up and drop off of packages. They allow the drivers to get in and out quickly with low step in height and easy access chairs and seatbelt. They also allow the driver to go into the package storage space from inside. They can search and pick packages easily. They may have a “sky light” to illuminate the package space. The packages are stored on shelves with random access capability. This allows mix the pickup with the delivery, like what UPS does at business locations.

2.4.5 Truck shipping rates and classes

TL costs are based on unit load, as long as within weight and volume capacity. Even if you have less than full truck, a full load is charged. The cost of LTL loads depends on the types of goods, the types of service, the distance, the origin and destination. The charges are often based on weight. However, the low-density goods can “cube out” a truck, meaning the loads reach the volumetric limit without reaching the weight limit. Therefore, the lighter goods will cost more per unit weight. The types of goods are divided into classes by the National Motor Freight Traffic Association (NMFTA). The cost per pound of freight goes higher as the class numbers go higher. The lowest class is 50 and is called Clean Freight. The highest class is 500 and includes those with low density or high value. The cost of TL is significantly lower than the cost of LTL per unit weight or volume. The differences can be more than double.

Between firms, the prices of regular services are often determined through contracts based on specific routes and other characteristics of services.

2.5 Intermodal transportation

The ISO containers are intermodal among ocean ships, waterway barges, train cars, and track chasses. This reduced the handling or double handling at the point of mode transition.

The trucking industry also adopted the ISO container interface on the 53, 48, and 40-foot trailers. These trailers, including the wheels, can be hoisted up by the cranes and loaded onto a flat car on a train. These are called trailer on flat car (TOFC). Each flat car can carry one trailer. In order to use the space better, a design that allow the separation of the chassis with the wheels and the cargo box. These road containers of 53, 48, and 40 feet long with ISO interface, can be hoisted up by the container cranes to stack two high on a train car. This doubles the capacity of the train.

2.6 Comparison of different modes

Table 6.1 summarizes the percentages of costs, the speed, and reach of different transportation modes, in order of increasing cost or the order of presentation above.

Table 6.1. Transportation mode, cost, speed, reach, and percentages of goods transported.

Mode	Cost / ton-mil	Speed	Reach	Ton	Ton-mil	Value
Water	Very low	Very slow	Drayage	8	11	5
Rail	Low	Slow	Drayage	10	30	3
Truck: TL	Medium	Fast < 500 miles	Door-to-door	60	35	69
Truck: LTL	High	Medium	Door-to-door	2		5
Package delivery	Very high	Fast	Door-to-door			
Air freight	Very high	Fast > 500 miles	Drayage			5
Pipelines	Low	Slow	If in-line	18	15	7

Due to these characteristics, the percentage of the transportation service value, ton, and ton-miles shared by these modes chosen by the users is shown in Figure 6.4 (Dobbins, Macgowan, Lipinski, 2007) [3]. The value is the value of the transportation service. Truck transported near 70% of the value, about 60% of tonnage, and about 35% in ton-miles. Therefore, trucking dominates U.S. freight transportation. It is also interesting to see that rail transportation service constitutes only 3% of the value, 10% of the tons, but over 30% in ton-miles. This is because the low travel cost on rail and the need for drayage makes it suitable for long-distance transportation. You may also notice that air has higher value than rail, < 1% in tons but appreciable ton-miles, mostly due to the long-distance effect. It is also interesting that parcel services have the third highest in value, near < 1% in tons, but some ton-miles.

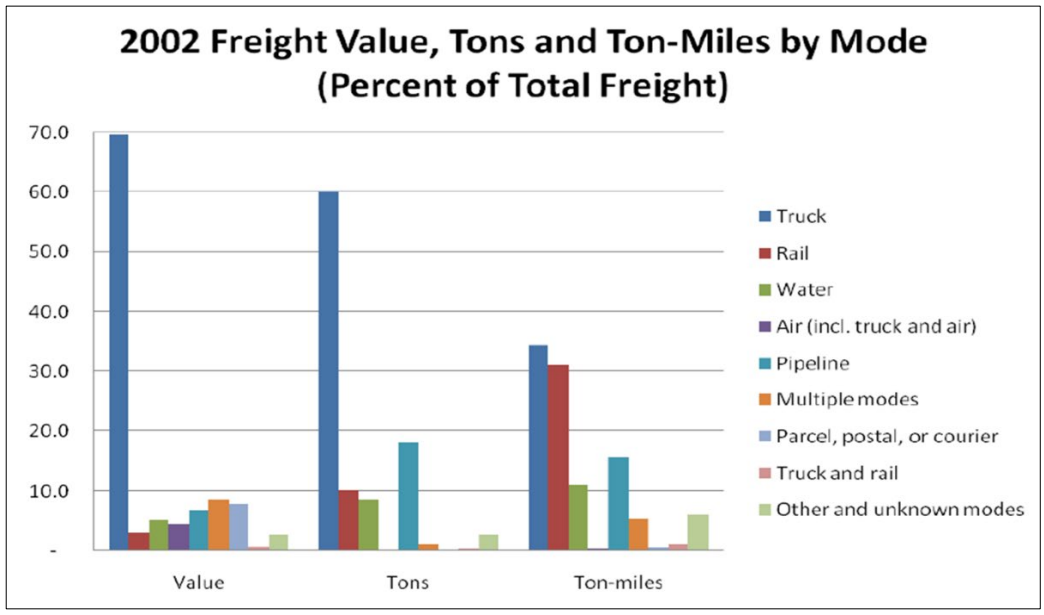


Figure 6.4. Value, tons, and ton-miles of different transportation modes.

Almost 98% of the truck tonnage is from TL. However, LTL receives about 7% of the truck revenue in 2007. This means that the average cost per ton in LTL is more than 3 times of average cost in TL.

2.7 Marginal cost of different mode of transportation

If we fix the distance at 500 miles, the marginal cost to transport one ton of goods varies significantly by mode. If we also fix the time requirement, the most important factor to select the mode is the scale. If a firm only sends and receives small volumes, Package is the only choice. In the medium volume, such as most manufacturing firms and logistics firms, the LTL, TL, and Rail are most relevant. The scale is the most important factor that determines mode selection. The relative transportation cost by mode is illustrated in Figure 6.5.

In Barilla's case, the weekly volume from Barilla to the distributor can enjoy a mixed item truckload. However, the volume from the distributor to the supermarkets may only be pallets or cases per delivery. Therefore, they must use LTL services at a higher cost per ton. The volume from the Barilla production facility to their two central distribution centers may enjoy the low cost of rail.

From parcel to LTL to TL to rail, the sizes of loads become bigger and bigger. To justify a bigger size, the demand must be higher in volume. Therefore, the economies of scale are very strong in transportation.

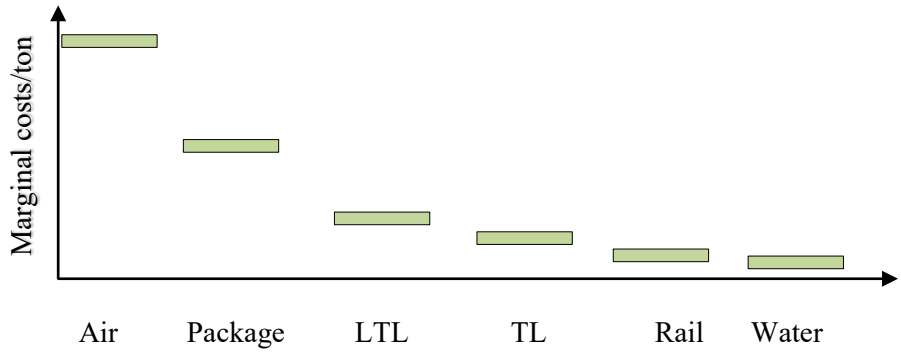


Figure 6.5. The relative range of the marginal costs of different types of transportation services.

3. STORAGE AND HANDLING IN THE SUPPLY CHAIN

In the supply chain, inventories are used to buffer against fluctuating demand and the delays in production and transportation. In the introduction, we showed a simplified supply chain for a calculator in Figure 2.1 in Chapter 1. The OEM receives inbound materials from its first-tier suppliers, who in turn receive their supplier from second-tier suppliers, and so on. The OEM distributes its supplies through retailers or online fulfillment centers, who in turn supply to the retailers or consumers. In every step, there are time delays in waiting, production or processing. Between every pair of nodes, there are delays in transportation. There is uncertainty at every node and in every arc in the supply chain.

The consumers have acceptable order fulfillment lead time, $LT_{acceptable}$, which is much shorter than the summation of the sum of the delays throughout the supply chain. For example, if I need a calculator, the maximum time I can wait may be two days, far less than what is needed to build the calculator from scratch. Therefore, the OEM must develop a strategy to keep inventory somewhere to satisfy the consumer needs within acceptable lead times.

3.1 Function of storage in a supply chain

Without inventory, when a customer wants something, the firms must start to make the product from raw materials. The total order fulfillment times can be weeks or months.

For the calculator, the OEM can work with the distributors and stores to carry finished calculators to meet the shopper's acceptable lead time, or $LT_{acceptable}$, of probably a day or two. If the calculator is out of stock, the shopper may go elsewhere or get a different brand. The store, the distributors, and OEMs would work to determine efficient replenishment cycles and quantities based on the demands and cost structures. They would keep the safety stock to balance between too much or too little. DC_2 is an online order fulfillment center. The online shopper may have longer order fulfillment lead times and lower cost with holding inventories. It can set its inventory accordingly. When the OEM works with its suppliers, it would consider similar situations.

3.2 Considerations in storage strategies

Inventory incurs costs due to the investment in the inventory, the storage, the potential for devaluation, damage, or even obsolescence. As the materials flows downstream in the supply chain,

1. The values of the goods increase due to more and more value-added operations to the materials.
2. The storage cost is the highest at the stores, because they are at high-cost commercial areas that require a customer-friendly environment with good lighting, air conditioning, and decorations.
3. The materials become less flexible to pool among different products. For example, a USB port is a USB port that can be used in any product that requires it. Once installed in a device, it is no longer available to other products.
4. They reduce the lead time to reach the customer, and therefore are more likely to satisfy customers acceptable lead time $LT_{acceptable}$.

A supply-chain professional must consider all these factors in formulating their inventory strategies to satisfy the customer demand timely, or $LT_{fulfill} \leq LT_{acceptable}$ with minimum cost.

3.3 Warehouses, distribution centers, and cross-docks in the supply chain

The materials are normally stored in warehouses and distribution centers. The two terms are often used interchangeably. A warehouse may be large space for storage, but the inbound and outbound flow volumes may be low. Distribution emphasizes high flow volumes. A cross-dock is an extreme case of a distribution facility: there is no dedicated storage space. The goods arrive on transporters are transferred to the outbound transporters without storage. Sometimes, the goods may be staged for a short time due to operational needs.

These facilities perform several functions.

1. Store or stage material as buffers.
2. Position inventory in strategic locations for cost or response time.
3. Break bulk or to convert high-volume, low-mix source to high-mix, low-volume destinations. For example, pallets in and items out.
4. Consolidate, or convert from high-mix, low-volume source to high-volume, low-mix destinations. For example, to batch products in a region to be transported to another region to reduce the transportation cost.
5. Add value. For example, the light assembly of adding customer-dependent parts to the unit, placing label or power plug with the right language or region.

3.4 Marginal cost associated with different handling units

Similar to transportation, the handling costs to receive, store, and dispatch goods depends on the handling units. The units can be item or piece, case, or shipping carton or pallet. An item can be a phone case you order online. A case is similar to a package you order or receive. In supply chain, a case is a package containing one or more of the same items. They are designed to be handled by a single person. It is often bigger for lighter items. A pallet holds many cases. It is designed to be handled by various pallet trucks.

The handling cost per unit weight depends on the handling units. A pallet of items requires many more touches and handling of cases, and a case of items requires many more touches and handling of items. Therefore, the marginal cost of handing items is much higher than handling cases and pallets.

Figure 6.6 shows the relative range of the marginal costs of different types of handling units. There are also mixed cases and mixed pallets. A mixed case contains different items. The purpose is to consolidate to move or to move before break the case to access the items. This is similar with mixed pallets. It can reduce part of the cost at the smaller units.

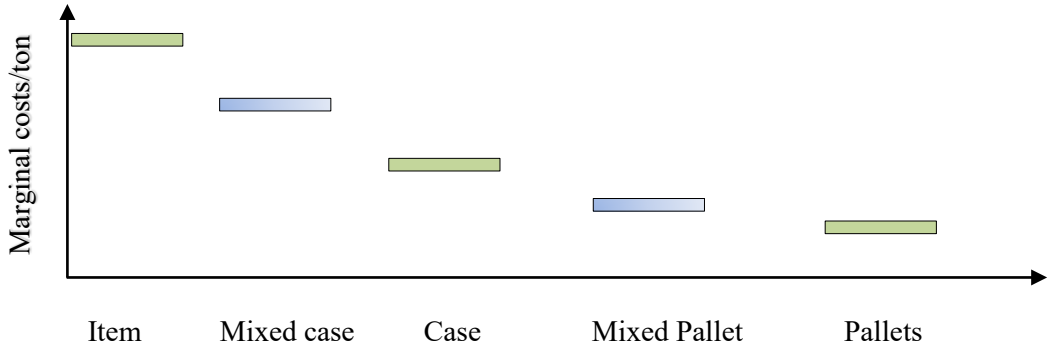


Figure 6.6. The relative range of the marginal costs of different types of handling.

From parcel item to case and to pallet, the sizes of loads become bigger and bigger. To justify a bigger size, the demand must be higher in volume. Therefore, the economies of scale are very strong in storage and handling in the storage.

4. LOGISTICS ADMINISTRATION

Administration refers to planning and executing transportation, storage, and handling in the logistics systems. The total administrative cost is much lower than transportation cost or inventory costs. Although the cost also increases over the years, it has expanded and improved the services and quality tremendously. The transportation management systems (TMSs), warehousing management systems (WMSs), together with enterprise resource planning systems (ERPs), manufacturing execution systems (MESs), and other systems have improved the supply-chain performance. They monitor the system, more and more toward real time, and use the data-driven solutions to support the decisions and control of the system. They help to improve the supply chain services with faster speed, better availability, lower inventories, and lower transportation costs.

In the initial Barilla case, the administration of transportation and inventories is traditional with person-to-person communications, manual key entry of orders based on the local and ad-hoc

information, and rough forecast to manage the logistics services. It led to a high level of inventories but also high stockouts and high transportation cost.

In the modern big-box retailer supply chain, represented in Figure 6.7, the retailer has visibility at the store, the vast network of distribution centers, and inventory in transit. The data are used to find best inventory levels at each location, through its WMS systems, and coordinate the in-house and outsourced transportation systems to deliver the goods to the right locations. In most situations, there is no need for stock review, ordering, and no needs for paperwork. Goods just arrive at the right locations under the administrative systems. The big box retailer can also make necessary data accessible by the suppliers for their production planning, and to coordinate the delivery from the suppliers to the distribution centers or to the stores directly.

In the modern supply chain, the administrative work is highly computerized. Many functions are automated, and optimized by the "experts." For each strategy, only one or a few experts develop tools that can be used system wide, or even make the tools available to others commercially or academically. The combination of information gathering, computing infrastructure, and expert knowledge, eliminates much of manual, repetitive, and localized work, and provides much better performance. Since the 2000s, artificial intelligence and machine learning created potential for even further improvements with less human input.

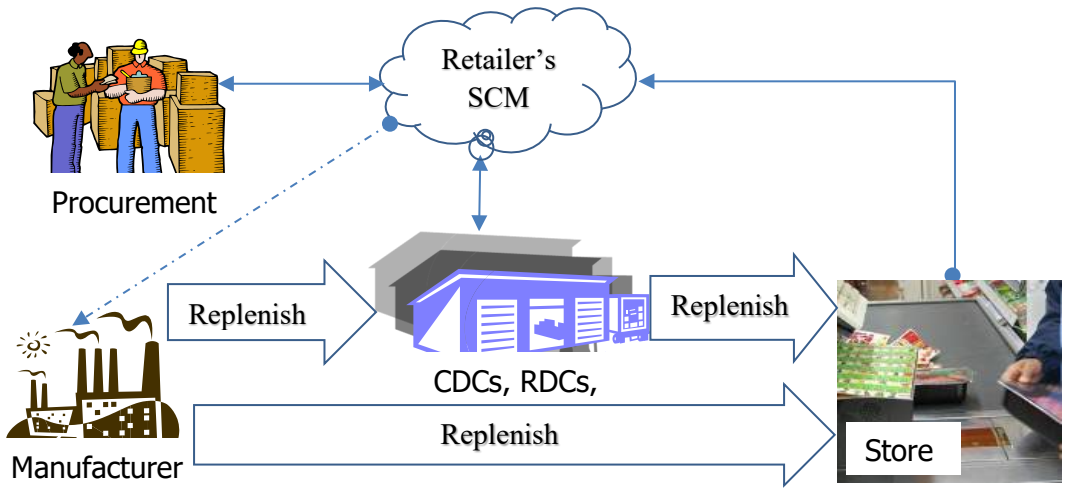


Figure 6.7. Simplified information and goods flow in a big-box retail chain.

5. LOGISTICS USERS AND SERVICE PROVIDERS

Firms start in certain sectors: fabrication, OEM, retail, or logistics service providers. Some firm grow to include other functions. Walmart has a large in-house fleet. Amazon started to have their own fleet, UPS started to provide maintenance services to HP, and so on. What are the issues in the service receivers and service providers?

5.1 Common carriers versus private fleets

The most expensive component of logistics services is transportation. A firm can develop their own private fleet or outsource to a common carrier, and possibly contract to a private fleet. Both common carriers and private fleets are subjects to government regulations.

A common carrier must provide services to any client. As a consumer, you may have used the common carrier package service providers such as UPS, FedEx, and DHL. There are many commercial freight service providers that provides transportation for businesses such as Schneider Logistics in (mostly TL), J.B. Hunt (dominate intermodal), ABF Freight Systems (LTL), UPS Freight (LTL), FedEx Freight (LTL). The common carriers publish shipping rates or tariffs through origin destination pairs, sometimes based on zip codes. In practice, they normally sign long-term contracts with their clients at much lower prices. In fact, even the client on the spot market may be able to negotiate a lower price. The use of a common carrier allows the firm to focus on its core business, and tap along the expertise, the scale, and other benefits the common carrier can provide. For example, an automobile seats supplier in Alabama needs one-way TL services going to the Kia plant in Georgia. If the seat supplier uses its own trucks, it would not have load on the return trip. In trucking lingo, this is called dead-heading. Some call it shipping air. However, a common carrier may have contract to ship something from near the Kia plant in Georgia to a client near the seat supplier in Alabama. This will reduce dead-heading.

Private fleets are to serve the firm itself mostly. They are allowed to provide limited services to other firms through contracts. Coca Cola and Walmart have significant private fleets. Coca Cola designs its vehicles to optimize their own operations and advertisement. Most importantly, a firm can have total control of their private fleet to line up priorities with the core mission. You may have seen the bright red Coca Cola cooled delivery trucks with small side openings. The small low-level access doors allow the driver to work with different packages of drinks conveniently. Many retailers also have their own private fleet, such as Target and Kroger. In fact, Walmart has one of the largest fleets in the country, private or common. One of the missions for most companies is growth. Adding a private fleet is to grow in scope. For Walmart, in addition to total control of the private fleet, Walmart also enjoys higher efficiency. Even though these firms own their own private fleet, they also use common carriers for certain services that do not complement their operations. For those, they hire common carriers or sometimes get services from smaller service providers.

5.2 Third- party logistics (3PL)

Most manufacturers, such as Texas Instruments, General Motors, Apple, and so on, nowadays outsource their entire logistics services.

Third-party logistics (3PL) provides both transportation and storage services. To differentiate, many call themselves fourth-party logistics (4PL), lead logistics providers or others. General Motors has large-scale and complex logistics services throughout the world. They hire a lead logistics provider who coordinates many services in broad scope around the world.

The benefit of the logistics providers is first their expertise. They would have the know-how in regulations, trade tariffs, relationships with various parties, and software platforms to optimize the entire chain. They also enjoy greater economies of scale. They can also use their large service

network to balance the seasonality, direction of flow, and availability of special equipment or applications. Therefore, they can often provide better services at lower cost.

The drawback of using logistics service providers is that most service providers serve multiple clients. Each client has its own contract with the 3PL. When there is contention in serving different clients, the 3PL must prioritize the services among different clients based on their relationships with each client and based on their own mission and resources. The prioritization scheme is 3PL's private information and not visible to the clients. This is especially true for the clients not considered high priority. In addition, if there is contention in serving different clients, such as different stores that sell Texas Instrument's calculators, the 3PL's priority may not align with the client, such as Texas Instruments.

Similar to a supplier-owned private fleet, the 3PL can also in-source services. They can provide value-added services to their clients. For example, UPS in Singapore provides certain maintenance services to HP. UPS can pick up the printers that need service and provide simple services inside the UPS facility. The collaboration benefits add revenue streams to UPS and at the same time provide faster services to HP's customers and with less transportation and handling.

The discussions in the chapters on collaboration in supply chain and principal-agent models can be applied to align the objectives and interests between service providers and receivers.

5.3 The factors in logistics system design

Regardless of logistics service providers or service demanders, large firms in the supply chain must have some knowledge of the logistics systems. For example, Texas Instruments must find which service provider is better aligned with their needs. Logistics services needs transportation and storage services. The former needs a fleet, and incurs the most cost. The later needs facilities, and can help to reduce the transportation cost.

The costs due to transportation and in the facilities can be shown in Figure 6.8. The cost in the facilities includes the cost of the facilities, the labor cost in the facilities, and the cost of inventories in all facilities. The cost of facilities, either rent or owned, increases with the number of facilities in step functions. The inventory costs normally increase with the number of facilities, because of the loss of economies of scale and other factors. The labor cost increases for similar reasons. One of the reasons to build facilities is to reduce transportation cost. Therefore, as the facilities increase, the cost may go down. However, with too many facilities, each trip delivery may not enjoy scale, such as TL load, and the transportation cost can go up. The total cost is the sum of all of the above, shown in Total.

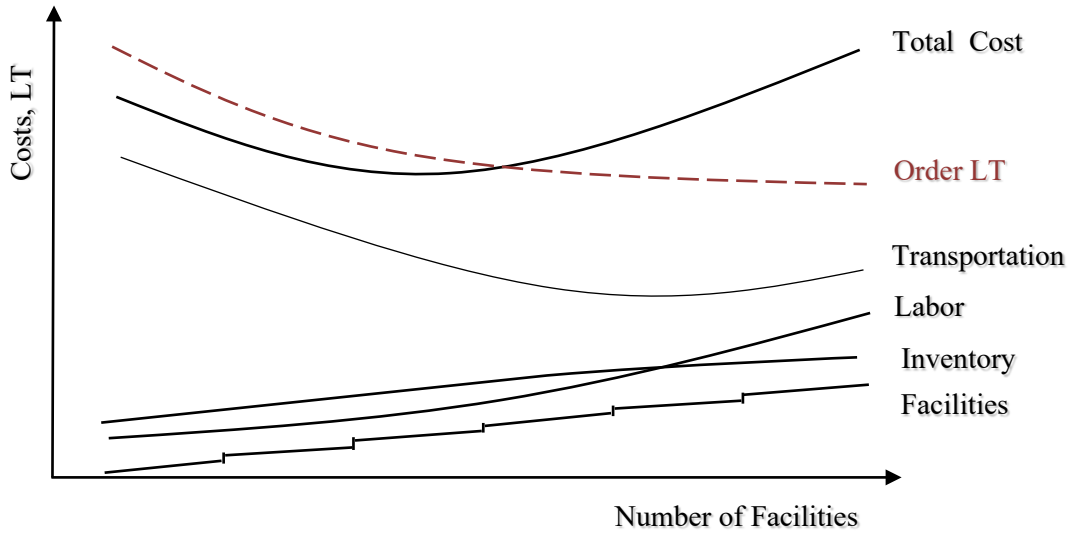


Figure 6.8. Various costs in logistics as a function of number of facilities.

One of the major benefits of more facilities is to potentially reduce the order fulfillment time, shown as the dashed line. As more and more facilities are added, it is possible to stock up inventories closer and closer to the client and therefore reduce the order fulfillment lead time.

5.4 Network effect in logistics

The network effect is the increased number of clients increasing the value of goods or services. Social media applications enjoy a huge network effect, because the more people on a particular platform, the greater value it is for all the users. This is true in logistics networks. For truckload services, more customer base can mean more opportunities to reduce dead-heading. For the 3PL, the more the coverage, the better it can serve its customers. If UPS does not cover a particular international address, it may not be able to provide you the service. It could collaborate with a partner in that country. However, the accountability and the traceability can pose challenges.

6. THE LOGISTICS IN THE BARILLA CASE

In the Barilla case, the product is commodities with low margins. The transportation and storage costs have significant impact on the profit. Barilla pays to transport the pastas to distributor's DCs. In order to reduce the transportation cost, Barilla offers a 2% to 3% discount for truckload orders. This is because TL costs much less per unit pasta. However, the bulky orders lead to slow responses. The results are excessive inventories for some SKUs and stockout to other SKUs. The cost of inventory is accounted for in the accounting system. However, the storage and the handling are normally not explicitly accounted for. The lack of information sharing and good administrative software led to poor visibility of orders, stock levels. The order processing times are normally long.

7. LOGISTICS AND HUMAN NEEDS

Human needs require goods to be delivered to where needed at the desired time and in the desired quantity, quality, and form. Modern logistics systems provide the transport and storage service globally at low cost. The containers enable the efficient global supply chain. In the developed economies, most people can find goods they need. Many goods and their components have travelled long distances, and were stored in various locations, before they reach the point of sale, such as fruits and low-cost merchandize from other continents, or vegetables from the other coast. Supply-chain professionals have made major contributions to this success.

However, there are still harder to reach locations or populations. For example, Amazon delivery does not cover many areas where the demand concentration is low. The cost of grocery delivery is higher than typical in-store shopping in supermarkets in most locations. In a food desert, there are additional challenges. In theory, the low-income household without car in the food desert can potentially enjoy grocery delivery through online ordering and delivery. However, such households may not have the credit and the means to order. The high cost of Amazon Fresh delivery also reduces the demand for the entire low-income neighborhood. As a result, the demand is sparse in the low-income neighborhood. The logistics cost will be even higher for Amazon to deliver. Therefore, Amazon Fresh is not a financially viable solution to satisfy the nutritional needs of the people in food desert.

Transportation is energy intensive. The U.S. Energy Information Administration (U.S. EIA) reported that the transportation sector consumed 37% of the total energy, with 91% being petroleum, in 2019. It generated about one-third of carbon emissions. The fresh air and ecosystems are external to the financial systems. We only pay for the transaction costs of the fuel, not the emission or the formation of the oil by Nature. The low cost of transportation does not align with human needs for fresh air and a sustainable ecosystem. At one point, the result of this distortion lead to the interesting situation that it was cheaper to butcher the chickens, transport to China for processing, and transport back to United States for consumption!

Today, many firms work to internalize the social and environmental measures into their financial systems. The good public image can even translate into financial values.

Logistics network such as roads, bridges, rail, ports, transporters and storage are part of vital infrastructure to supply for human needs. However, the cost driven infrastructure is often vulnerable to attacks from nature or adversaries. The failure in logistics network can hinder the supply to human needs severely, as we have seen over the years. Therefore, the supply chain and infrastructure resilience are hot topics nowadays.

7. ECONOMIES OF SCALE IN SUPPLY CHAIN

In Chapter 4 Competition and Chapter 6 Logistics, it was clear that in any competition, the costs play an important role. In Competition, the assumption was that the average cost does not vary with quantity. In practice, the costs can change significantly with supply quantity as shown in the chapter on Logistics. Economies of scale refers to the effect that the average cost of production decreases as the production quantity increases. This is also called the scale effect. The scale effect can be very strong in certain types of supply-chain systems and is an important strategy in supply chains. It is part of the motivation for a firm to grow, merge, and acquire. In Logistics and Collaboration in Supply Chain, we discussed that a firm can outsource its logistics services to 3PLs to enjoy economies of scale.

There are also diseconomies of scale in which the average cost increases with quantity, such as when the supply reaches the system capacity. Larger quantities may lead to greater complexity, additional work contents that need more management, coordination, overtime pay, extra capacity, and so forth. The dry cleaners, maintenance shops, and alternation shops tend to stay small because of this. However, information technology and artificial intelligence may offer opportunities to change that.

1. ECONOMIES OF SCALE IN ECONOMICS AND IN THE SUPPLY CHAIN

1.1 In economics

The economies of scale model is defined as when the production quantity increases, the cost decreases. Mathematically, this can be expressed using production quantity and a few measures below.

Definitions

q Production quantity

$TC(q)$ Total cost as a function of production quantity

$AC(q)$ Average cost as a function of production quantity

$MC(q)$ Marginal cost, incremental cost to product q + 1st unit, as a function of the production quantity.

With these definitions, the relationship between these cost functions is

$$AC(q) = \frac{TC(q)}{q}$$

For the marginal cost in discrete space, we can consider adding the current unit or another unit. Let us use the latter. We have

$$MC(q) = \frac{\Delta TC(q)}{\Delta q} = \frac{TC(q+1) - TC(q)}{q+1-q} = TC(q+1) - TC(q)$$

Graphically, these are shown in Figure 7.1.

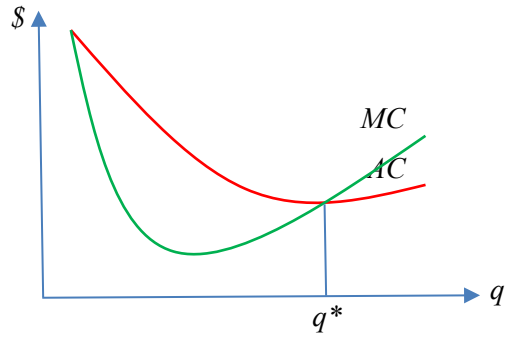


Figure 7.1. Average cost and marginal cost.

When $q < q^*$, the average cost goes down when production quantity increases; there are economies of scale. At $q = q^*$, *the* $AC(q^*) = MC(q^*)$, *and* $\frac{dAC(q^*)}{dq} = 0$. When $q > q^*$, the average cost goes up when production quantity increases; there are diseconomies of scale.

A common question about a product or service is, "does something scale up?" This can include the effect of the economies of scale.

1.2 Marginal and average costs in the supply chain

In supply chains, the economies of scale are results of some specific factors in the supply chain that supply-chain professionals must consider. Many of these can be quantified in concrete expressions or good approximations. Here are some factors that can be used to approximate the total cost functions.

1. Internal factors contribute to the economies of scale
    a. Shared resources: higher production quantity allows more units to share the cost of investment or equipment.
    b. Technological or system design: systems with higher production rates normally incur higher initial cost but lower marginal costs. A higher production quantity may justify the higher initial cost with lower marginal cost through mechanization, automation, IT, or a different system structure, such as a flow line.
    c. Batching: higher supply quantity leads to larger economic batches to reduce the cost per unit due to sharing of resources or setups, such as moving pallets instead of cases or changing to different product types.

   d. Shared overhead: marketing, management, facilities.
2. External sources of scale effect
   a. Unit load: In the supply chain, there are well-established unit loads for scheduling, handling, and transportation such as eight-hour work shifts and pallet loads or truckloads. When the quantity is less than a unit load, the cost incurred is similar to the full-unit load, such as if you transport goods in a truck.
   b. Quantity discount in raw materials: higher production requires more raw materials and services and enjoys quantity discount.

Most of these factors you have already learned. You can use them to provide approximations to enhance communication to those outside of the supply chain.

1.3 Diseconomies of scale

Theoretically, when the average cost starts to increase, or when the marginal cost increases to pass the average cost, diseconomies of scale emerge. In what situation can the marginal cost increase that fast? Here are a few factors that can occur in the supply chain.

1. When the production exceeds normal capacity, there can be overtime payments to the personnel, outsourcing the excess production, and production using less efficient facilities. An example will be given later.
2. When the supply is for markets with higher costs, such as a market with a lower demand density or in a foreign country, there is additional cost in transportation, inventories, tariffs, and additional services. Examples will be given in later sections.
3. When the investment cost is low, coupled with increasing marginal cost. Such industry tends to stay small, such as dry cleaners, salons, alterations, repair shops, specialized consulting firms, freelance coders, and so forth. For example, a software developer uses the number of lines of code as the measure of production quantity; the complexity of software grows as the number of lines increases due to more and more interdependencies between the lines. Another example is the number of people who work in a firm. As the number of people grows, the organization becomes more complicated. These are internal measures. If the software developer uses the number of clients it serves, more common from the supply perspective, the result may be different.
4. When there is scarcity in the external resources such as raw materials, public infrastructure such as roads, and the water supply. The marginal cost can increase rapidly.

In the supply chain, there are factors that contribute to the economies of scale significantly. We will expand on a few of the ones listed above.

2. Scale Effect due to Lower Shared Investment Cost

For a given facility or a piece of equipment, there is an initial cost. For each unit produced, there is an incremental cost. We can first explore the cost characteristics of a given system.

Definitions

I Investment cost or all costs independent of production quantity. This can be net present value or annual equivalent of investment cost. If the annual equivalent cost is used, the annual production quantity should be used.

mc Marginal or incremental cost to produce one unit. If marginal cost in a supply chain is normally a constant within its designed normal capacity, then going beyond the normal capacity, a system might use overtime or other measures to increase the production at higher marginal cost.

C Capacity.

Consider a simple case in which the initial investment cost is *I*, including all other fixed costs. Its designed normal capacity is C_0 units. The marginal cost within this capacity is a constant, mc_0. A firm can produce more than the designed capacity by running overtime. The marginal cost in the overtime is higher at mc_1. We can simplify the problem to find the expressions for marginal, total, and average costs.

$$MC(q) = \begin{cases} mc_0, & 0 < q < C_0 \\ mc_1, & C_0 < q < C_1 \end{cases}$$

The total cost is

$$TC(q) = \begin{cases} I + mc_0 q, & 0 < q < C_0 \\ TC(C_0) + mc_1(q - C_0), & C_0 \le q \le C_1 \end{cases}$$

The total cost under the designed capacity is the fixed cost plus the marginal cost multiplied by the production quantity. The total cost with overtime is the total cost at the designed capacity plus the marginal cost during overtime multiplied the production quantity during overtime, or the difference between the total quantity and the normal capacity.

The average cost is

$$AC(q) = \begin{cases} mc_0 + \dfrac{I}{q}, & 0 < q < C_0 \\ \dfrac{TC(C_0) + mc_1(q - C_0)}{q}, & C_0 < q \le C_1 \end{cases}$$

The marginal costs and average cost are shown Figure 7.2.

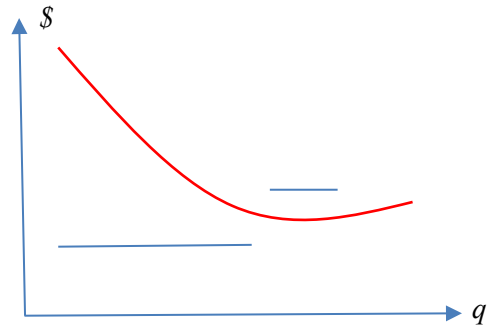


Figure 7.2. The marginal and average cost in regular time and overtime of a facility.

Although this very simple form omits a lot of complex situations, it provides great insights:

1. The average cost is a decreasing function up to capacity. The higher the quantity, the lower the cost per unit.
2. Beyond the design capacity, the marginal cost is much higher. The rate of decrease in average cost initially reduces and at some point, can stop decreasing and start to increase. That is the point that separates the economies of scale or diseconomies of scale.

Recall in the economies of competition, the assumption was that the marginal cost is a constant. Now you can see this is a good assumption in practical situations within the designed capacity range.

In the United States, supply-chain firms rarely run full design capacity. Some do, some do sometimes, and others often run overtime. If the demand consistently exceeds the design capacity, the firm may consider additional investment to increase the capacity to avoid higher marginal cost. For many industries, it is difficult to reach the point of diseconomies of scale at a macro-level, such as those continue to expand to enjoy lower costs.

Example

In February 2020, the demand in the United States for N95 masks surged due to the pandemic. A multinational firm has its main factory is in China where they produce N95s but mostly KN95s, which are not approved by the FDA. Due to the rapid increase in demand in China and worldwide, trade tension, and FDA approval, the firm must increase its production in the United States. The current production system in United States is 50,000 units/year. The investment or fixed cost is $500,000 per year. The marginal cost is $50/unit. The average demand is 35,000 per year with some variability. The demand in the current year and possibly next will be over 200,000.

The first thing to do is to ramp up the production to its maximum capacity of 50,000. Above the capacity, the firm can add some work hours, hire back the laid-off workers, and take other measures to reach the capacity of 80,000. The marginal cost is increased to $60 per unit. To increase further, the marginal cost is $100/unit up to a capacity of 100,000.

Questions:

1. What are the marginal costs for demand from a quantity of 1 to 100,000?
2. What are the total cost functions at different marginal costs?
3. What is the average cost?
4. Please plot these costs, compare to the general curve in Figure 7.1, and identify the characteristics.

Solutions

1. The marginal cost is

$$MC(q) = \begin{cases} 50, & 0 < q < 50{,}000 \\ 60, & 50{,}000 < q < 80{,}000 \\ 100, & 80{,}000 < q < 100{,}000 \end{cases}$$

2. The total cost is

$$TC(q) = \begin{cases} 500{,}000 + 50q, & 0 < q \leq 50{,}000 \\ TC(50{,}000) + 60(q - 50{,}000), & 50{,}000 < q < 80{,}000 \\ TC(80{,}000) + 100(q - 80{,}000), & 80{,}000 < q < 100{,}000 \end{cases}$$

Where,

$TC(50{,}000) = 500{,}000 + 50 * 50{,}000 = 3{,}000{,}000$, and

$TC(80{,}000) = 3{,}000{,}000 + 60(80{,}000 - 50{,}000) = 48{,}000{,}000.$

We can simplify the total cost to

$$TC(q) = \begin{cases} 500{,}000 + 50q, & 0 < q < 50{,}000 \\ 60q, & 50{,}000 < q < 80{,}000 \\ -3{,}200{,}000 + 100q, & 80{,}000 < q < 100{,}000 \end{cases}$$

3. The average cost

$$AC(q) = \begin{cases} 50 + \frac{500{,}000}{q}, & 0 < q < 50{,}000 \\ 60, & 50{,}000 < q < 80{,}000 \\ 100 - \frac{3{,}200{,}000}{q}, & 80{,}000 < q < 100{,}000 \end{cases}$$

4. The plot is shown in Figure 7.3.
    a. The marginal costs are short and dashed lines. Its scale is on the left. It is a step functions that remain constant and jumps at the critical capacity constraints. This is common in most supplies in production or transportation within one facility.
    b. The solid line is the total cost. Its scale is to the right. It increases with production quantity and increase the rates at two thresholds at q = 50,000 and 80,000.
    c. The average cost is a convex curve. When q < 50,000, it decreases with quantity. From q = 50,000 to 80,000, the average cost remains constant. This is coincidental. Above 80,000, the average cost bent up. Between q=50,000 and 80,000, $MC = AC$. This is also coincidental. They may be different.
    d. Beyond 80,000, $MC > AC$.

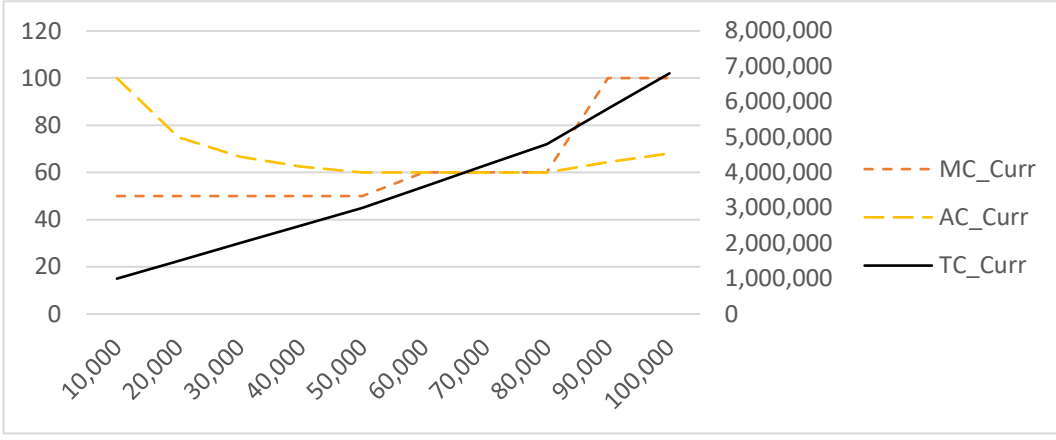


Figure 7.3. The marginal cost, average cost, and total cost near the break-even point.

3. ECONOMIES OF SCALE DUE TO TECHNOLOGY AND SYSTEM DESIGN

When the production quantity is low, firms may invest in small-scale systems with lower investment cost, low throughput and capacity, but higher marginal cost. If the demand increases beyond a certain point, firms can investment more for mechanized or automated processes with higher investment cost, higher throughput, but lower marginal cost. There is a significant marginal cost jump at the point of system change.

For example, a simple printer/scanner/copier has a low investment cost. They are slow, highly manual, and with high marginal cost. They may be sufficient for small operations. If the demand becomes high, the firm can invest in a fast and highly automated system with low marginal cost. The same is normally true for the facilities used for many commercial products or services such

as food processing, plastic molding, electronic assembly, car assembly, operating room warehouses, or freight transporters.

Consider a simplified case with two alternatives. The baseline system has lower investment cost, higher marginal cost, and lower production capacity. The alternative has higher investment cost, lower marginal cost and higher production capacity. The marginal costs and average costs of both the baseline system and alternative system are shown in Figure 7.4. The dashed lines are for the baseline system and the solid line for the alternative system. The green lines are for the lower-cost alternative at a certain supply range considering both alternatives.

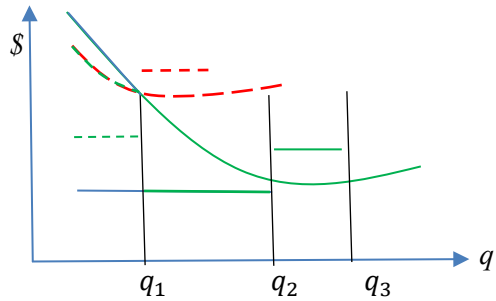


Figure 7.4. The marginal costs and average costs for baseline and alternative systems at different production quantities.

The marginal costs for each alternative remain constant within its designed capacity for each alternative. Therefore, they are shown as horizontal lines. However, when the production is higher than the designed capacity, the marginal cost can increase, such as running overtime. During overtime, the marginal cost can also be constant until it reaches its capacity. If the firm wants to produce more, extreme measures may be needed that can lead to even higher marginal cost, or even additional investment cost for higher capacity systems. The average costs are shown as curved lines. At each quantity range, one of alternatives yields the minimum average cost. In Figure 7.4, the average cost of the baseline system is lower than alternative system when $q < q_1$, and the opposite is true when $q > q_1$. When $q > q_2$, the demand exceeds the capacity of the second system. The marginal cost increases. The green lines show the lower-cost alternative at different demand ranges.

Example

In the N95 example, the capacity in the firm's current system is too low even at its maximum capacity at high cost. The firm is considering renovating an unused plant with the most modern facilities. If the new facility can reach the production capacity in 6 months, the demand should still be high. With the lessons learned during the pandemic, the firm also considers ramping up domestic production for the long term.

The investment cost will be $2 million per year. The marginal cost is only $20 per unit, with a capacity up to 200,000 units per year. With some added investment cost and at higher marginal cost, it can reach 250,000 units per year.

Questions

1. What are the marginal costs, total cost, and average cost for supply from 1 to 250,000 units per year?
2. Please plot the curves and compare the curves with the current system.
3. What should the firm do with the new and current facility? What are the total cost functions at different marginal costs?

Solutions

1. The marginal cost is

$$MC(q) = \begin{cases} 20, & 0 < q < 200{,}000 \\ 50, & 200{,}000 < q < 250{,}000 \end{cases}$$

The total cost is

$$TC(q) = \begin{cases} 2{,}000{,}000 + 20q, & 0 < q < 200{,}000 \\ -4{,}000{,}000 + 50q, & 200{,}000 < q < 250{,}000 \end{cases}$$

The average cost is

$$AC(q) = \begin{cases} \frac{2{,}000{,}000}{q} + 20, & 0 < q < 200{,}000 \\ -\frac{4{,}000{,}000}{q} + 50, & 200{,}000 < q < 250{,}000 \end{cases}$$

2. The curves for the new system and current systems are shown in Figure 7.5.

The three shorter and dimmed curves are for the current system. The longer curves are for the new system.

When $q < 50{,}000$, the current system incurs lower average cost per unit. If $q > 50{,}000$, The new system incurs lower average cost.

Since the demand is so much higher than the capacity within 1 to 2 years even with the new capacity, the firm may want to invest in the new system. The firm may have to run both systems after the new system is ready even though the current system incurs a lower cost. There is not sufficient capacity.

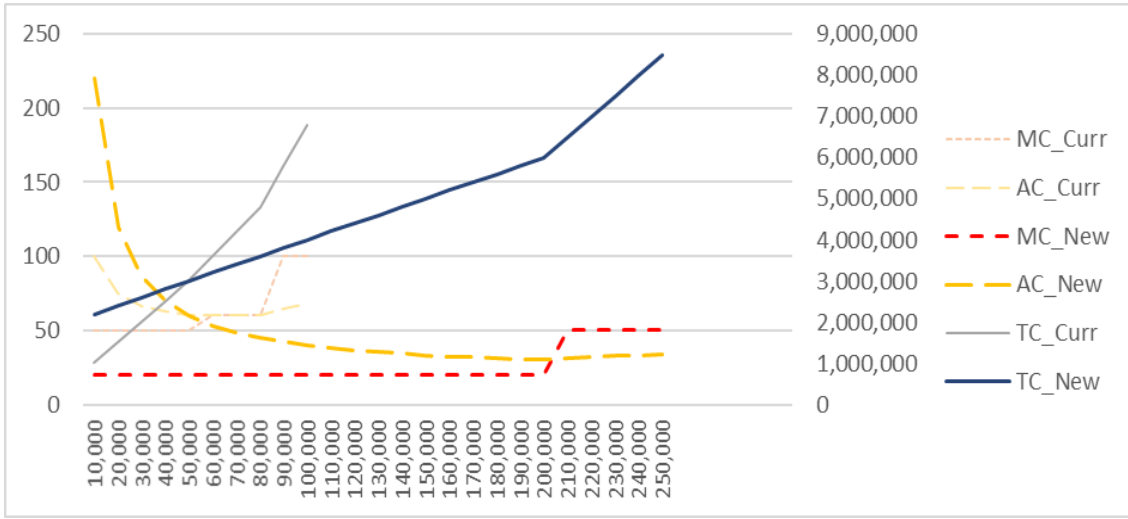


Figure 7.5. The marginal, average, and total costs of two systems.

Example

A firm plan to make gears for wind turbines. An engineer found a traditional option and a computerized option. The investment cost for the traditional option is $60,000 per year, with incremental cost of $60 per gear. The annual maximum capacity is 1,000 gears per year. If the demand exceeds the capacity, the firm must adjust production with a cost of $10 extra each gear. Up to 1,200. Beyond that, the firm may have to outsource or through other ways with much higher cost at estimated marginal cost of $120. The investment cost for the computerized option is $90,000. The incremental cost is $35 with maximum capacity of 2,000. The demand is estimated to be uniformly distributed between 700 and 1,200 gears per year. In addition, the demand is expected to increase 10% per year over next five years.

Questions

1. What are the total, average, and marginal costs for each option?
2. Please plot the total and average costs for both options.
3. Please find the break-even quantity between two options.
4. If your manager asks you to select one from the two, which one would you pick?

Solutions

1. Please find and plot the total cost, average cost, and marginal cost as a function of the production for the traditional and computerized systems.

For the traditional system, the marginal cost is

$$MC_t(\mathrm{q}) = \begin{cases} 60, & 0 < q < 1000 \\ 70, & 1000 < q < 1200 \end{cases}$$

$$TC_t(q) = \begin{cases} 60000 + 60q, & 0 < q < 1000 \\ 120000 + 70(q - 1000), & 0 < q < 1200 \end{cases}$$

$$AC_t(q) = \begin{cases} \frac{60000}{q} + 60, & 0 < q < 1000 \\ \frac{120000+70(q-1000)}{q}, & 0 < q \leq 1200 \end{cases}$$

For the computerized system, the designed capacity is sufficient. If we only consider the designed capacity, we have

$$MC_c(q) = 35,\ 0 < q < 2000$$

$$TC_c(q) = 90000 + 35q,\ 0 < q < 2000$$

$$AC_c(q) = \frac{90000}{q} + 35,\ 0 < q \leq 2000$$

2. The plots for total costs and average costs of both options are show in Figure 7.6. The increasing lines are for total costs. Both pairs intersect at the same breakeven quantity.

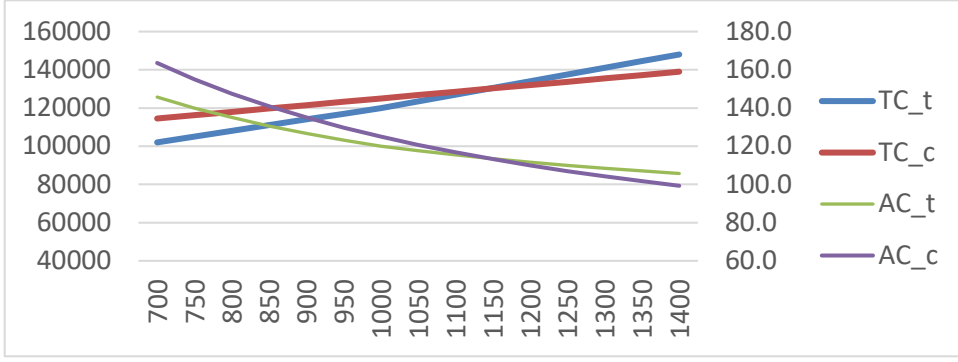


Figure 7.6. Total costs and average cost for the example.

3. The break-even points between two options can be found using total cost or average costs.

$$120000 + 70(q - 1000) = 90000 + 35q$$

$q = 1142.85$

We can round it up to 1143.

4. The average demand is 950, below the break-even point. The cost when the demand exceeds capacity is not given. If the added cost is less than $10 per unit. The expected cost for the traditional option is lower for the first two years. However, the cost will be higher from the third year on. If the firm plans to produce such gears for longer than two years, the computerized machine is recommended. In addition, the computerized machine may also have benefits if the firm decided to make other types of gears.

4. ECONOMIES OF SCALE DUE TO BATCHING

People batch to save on something. When you bake cookies, you may want to bake ten instead of the three you would eat fresh. When you need a bottle of water, you pick up more than one bottle to save another trip when you need water again. The trip incurs fixed cost such as time to travel, fuel, wear and tear on the automobile, and possibly parking.

4.1 Batching to unit load

In the supply chain, there are well-designed infrastructure components to move, store, and manage unit loads. The more common unit loads in increasing sizes are:

1. Production batch. Many things are produced in batches in a certain mixture, temperature, pressure, or other conditions for a certain amount of time, regardless of the quantity.
2. Cases, totes, or shipping cartons. These are cubical containers made of corrugated cardboards or plastics designed for one human operator to pack, move, and stack .
3. Pallets, skids, crates, or wire baskets. These are holders made of wood, plastic, or steel for a wide variety of pallet-handling equipment such as forklifts to stack and move many cases or other items.
4. Containers. There are ISO standard containers made of steel to be handled by container-handling equipment such as forklifts and cranes to move many pallets, cases, or other items. The ISO containers are 20 or 40 feet long, 8 feet wide, and 8 feet or 9.5 feet long, with weight limitations. Note that the weight limit for a 40-foot container is not double the weight for the 20-foot one. There is entire infrastructure on ocean liners, train carts, and truck trailers and chasses designed to move and store these containers.
5. Trucks. In the United States, common truck lengths are 53, 48, or 40 feet long. Each truck requires the fixed cost of one truck and one driver. There are also intermodal containers of these sizes to be handled between trucks and rail, that can be handled with the same cranes.

Large industries have developed to support different types of unit loads in terms of processing, packing, storage, moving, transporting, and protecting. However, small operations that do not meet the unit loads quantity will cost more per unit. Here are some examples.

You want to bake 10 cookies at home. The oven capacity is 30. However, the cost baking process for 10 and 30 cookies are about the same. This is true for brewing beer, chemical processes, heat treatment, and many other batch productions.

An order-picker travels to a location to pick up a textbook for an order, and then returns for the next order. There may be another order requesting the same textbook. The operator could pick up both textbooks without difficulty. In fact, there may be 10 orders in the morning, all request the same textbook. An order picker can use a tote with a 10-textbook capacity to pick up all 10 in the same trip. This will reduce to cost significantly. Furthermore, the order-picker can also push a cart with 10 totes to pick up 10 different textbooks in one trip.

The above example is also true for transportation. The cost per case would be lower if the quantity can reach unit load of pallets, a container, or a truck.

You can find examples in office work, health systems, and many other systems. Therefore, a system with production less than unit load will lead to higher cost due to wasted capacity. In transportation, some call this "shipping air."

In unit handling, regardless of whether a case, pallet, container, or truck is full, the fixed unit load cost dominates the entire cost. The total cost is shown in Figure 7.7. The unit cost is always lowest at the multiples of full units.

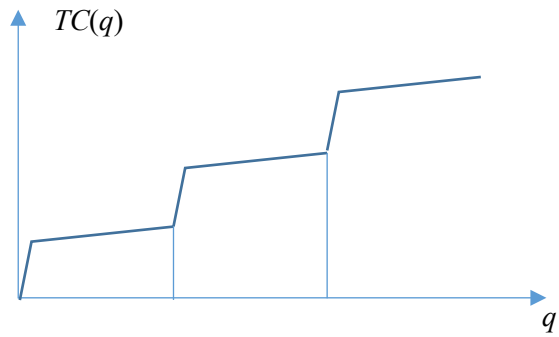


Figure 7.7. The cost characteristics associated with unit loads or unit batches.

Example

A firm supplies 10 tons of food in 20 pallets to a distribution center every week. A truckload (TL) costs $1,500 from origin to destination. A truck can take up to 20 tons for up to 30 pallets. A truck can be cubed out for bulky items or weight out for heavy loads. Since the amount is less than a truckload, you may also use a less than truck load (LTL) carrier. The LTL charges (after negotiation) $150 per ton.

Question

1. Which service should you select?
2. What if your demand increases 50% every year?

Solutions

1. The total cost for both TL and LTL is the same at $1500. From a cost perspective, they are the same. However, the LTL carriers normally consolidate loads, which leads to more stops, load and unloads of pallets, and so on, which leads to longer time, more handling, and more likelihood for errors. Therefore, TL is preferred.
2. If TL is used, the transportation cost remains the same for next two years. In the third year, the demand will be 20 tons and 40 pallets. They can negotiate with the distribution center, such as twice a week. This will help the firm to reduce its own inventories.

4.2 The dimension of batching

In a supply chain, batching can be in many dimensions.

1. Time. You can batch over time. If the demand is half of a TL, two weeks would become a TL. There will be significant savings in transportation of two LTL than a single TL. However, batching over time has its drawbacks. First, the service time or lead time becomes longer. Imagine if UPS changed delivery to every other day. If there is inventory involved, both the supplier and client have to carry more inventories. These contribute to the diseconomies of scale.
2. Different products or services. In production, different products can be designed so that they can be processed in the same batch. Equipment can be designed to be able to process different products in the same batch. In transportation, goods with different origin and destination can share the transporter during the same leg in the transportation. Amazon saved a lot by outsourcing many deliveries to USPS, which delivers to each location every day anyway.
3. Different customers. A firm can expand its customer base so that their demand can share the same resources.

Each one of these dimensions of batching has drawbacks and can lead to diseconomies of scale.

5. ECONOMIC ORDER QUANTITY (EOQ) AND LOT-SIZING

Consider buying water again. Although you may want to pick up more than one bottle for the immediate needs, you would not buy 100 bottles considering the storage space and cash flow. This is true in supply-chain firms in production and purchasing.

In production systems, it is necessary to change over between batches of different products. A changeover takes time, labor, and possibly materials and tools. The larger the batch, the more identical products can share the same changeover cost, and so reduce the overall cost, similar to the general economies of scale. The difference is that the lot-sizing problem is repeating.

In purchasing, generating orders, shipping, and packaging add cost, similar to when you order something online.

5.1 EOQ model

EOQ provides an excellent approximate quantity to determine optimum lot sizes. Although the model and calculations are simple, many often have misconceptions. You may have learned this elsewhere, but we will repeat it here.

Assumptions

1. A product analyzed independent of other products and services
2. The demand is continuous with constant rate D per unit time
3. Replenishment is instantaneous
4. No shortage is allowed (or the shortage cost is prohibitive)

Definitions

K Setup cost or change over cost per setup or per changeover. This can be the transportation cost each trip, adjustment needed for a particular product, paperwork needed to issue a purchasing order, and so on. We assume that it is not a function of purchasing quantity.

H Holding cost per unit per unit time. This can include storage cost and time value of money.

c The cost of each unit.

Q Production or order quantity per changeover or per setup. This is a decision variable. Here, we used a capital Q to differentiate from a lower-case q earlier for the total production quantity per unit time or for the life of a facility. In inventory problems, increasing Q will lead to less frequent production, less changeover but higher inventories. The production per unit time is to satisfy the demand, D.

Note that K, h, and c are not functions of Q. Many of these assumptions can be relaxed to accommodate more complex problems.

Under all the assumptions, the inventory level over time can be shown in Figure 7.8.

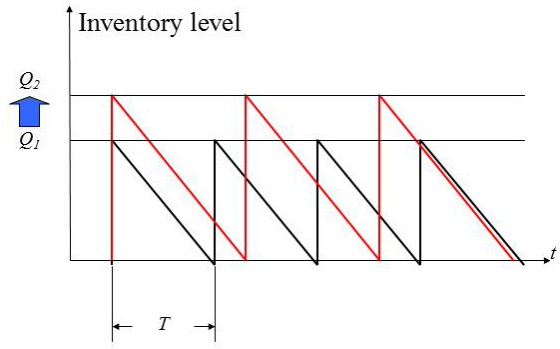


Figure 7.8. Inventory level changes over time at different order quantities.

The slope is the demand rate. The lowest inventory level is zero, because no shortage is allowed. The highest inventory level is reached just after replenishment. The objective is to minimize the total cost per unit time, which can be written as the sum of setup cost, holding cost, and purchasing price per unit time as

$$min\, C\,(Q)/unit\ time\ = K\frac{D}{Q} + h\frac{Q}{2} + Dc$$

The first term is the total cost due to setups, or the setup cost times number of setups. The second term is the carrying cost. It is the holding cost per unit time times the average inventory level, which is half of order quantity Q. The last term is the prices you pay for the goods. The last term is normally the highest in the total cost. However, it is independent of optimum order quantity as long as the cost is not a function of order quantity, such as in quantity discount.

The decision variable is the order quantity. Only the first two terms are influenced by the order quantity and can be considered inventory-related cost.

The larger the order or lot size, Q, the more pieces will share the setup or ordering cost and enjoy the economies of scale. However, the more you order, the more you have to carry, and that requires capital, occupies space, and subjects to risk (loss, damage, price drop, obsolescence). The ordering cost, carrying cost, and total cost as a function Q are shown in Figure 7.9.

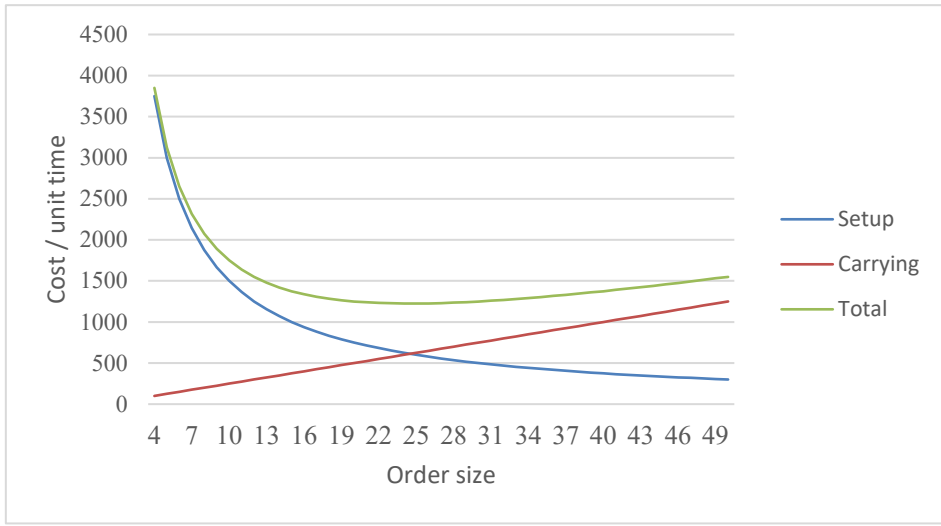


Figure 7.9. The trends in inventory-related costs as a function of Q.

If you apply first-order condition to the total cost per unit time model above, you can find the optimum order quantity, commonly refer to as the Economic Order Quantity (EOQ),

$$Q^* = EOQ = \sqrt{\frac{2KD}{h}}$$

Another useful measure of the quantity is how long does the inventory last between orders, or inventory cycle time, T, or in industry, people call it days of supply or DOS.

$$T^* = \sqrt{\frac{2K}{Dh}}$$

Example

The demand in an auto repair shop for a popular filter is 200 per week. The ordering cost is $100 per order, regardless of order size. The cost of a filter is $16. The shop set holding costs at $0.04/week. What are EOQ, the order cycle time, and cost at EOQ?

Assumption check

1. An auto part store may order different models of the filters from its suppliers. If they are ordered together, the independent assumption is violated and can have impact.
2. The demand for the filers is discrete, and uniform, some days more, some days less, and sometimes a single customer may get several. The continuous and constant rate assumptions are violated.
3. The replenishment takes place after receiving and restocking. The time from the restocking of the first item to last should be short compared to the order-cycle times. Therefore, instantaneous replenishment assumption is good.
4. The no stockout assumption may not hold, because the demand of 200 is uncertain. The stockout cost may be high for such commodities: a stockout is most likely a lost sale with lost profit and loss of good will, and the customer may not come back.
5. The unit cost and unit holding cost should not be a function of order quantity within a small range. The setup cost is unlikely to be influenced by the order quantity. Those assumptions hold.

Solution

Even though assumptions 1 and 2 are violated, we can still calculate the EOQ. We just need to know the result is an approximation.

$$EOQ = \sqrt{\frac{2*100*200}{0.04}} = 1000 \text{ filters}$$

$$T = \frac{EOQ}{D} = \frac{1000}{200} = 5 \text{ weeks}$$

$$C(1000) = 100\frac{200}{1000} + 0.04\frac{1000}{2} + 200 * 16 = 40 + 3200 = \$3240 \text{ per week}$$

Both the costs of setup and carrying are $20 per week, respectively. The sum of the two of $40 is related to inventory policy. The $3,200 purchasing cost for the filters will be recovered once they are sold.

The inventory- related costs at other order quantities from 200 to 1800 are shown in the table below.

	Setup	Holding	Purchasing	Total
200	100.0	4	3200	104.0
400	50.0	8	3200	58.0
600	33.3	12	3200	45.3
800	25.0	16	3200	41.0
1000	20.0	20	3200	40.0
1200	16.7	24	3200	40.7
1400	14.3	28	3200	42.3
1600	12.5	32	3200	44.5
1800	11.1	36	3200	47.1

Note even at 800 or 1200, the costs are only increased less or equal to 2.5%! Therefore, do not take hold the optimum Q with too many significant digits. Considering the violations in the assumptions and flatness in the curve around optimum Q, two or three digits should be sufficient in most practical situations.

The characteristics of EOQ:

1. Proportional to the square root of the demand. Therefore, batching is related to the square root of economies of scale.
2. Proportional to the square root of the setup cost. If you can reduce the setup cost, you can reduce the inventory level, in a lesser way.
3. Proportional to the inverse of the square root of holding cost. Does this mean a higher holding cost is better? You can jack up holding cost by using fancier facilities. How can that make sense?
4. The inventory-related cost, the sum of setup, and carrying costs are normally much lower than purchasing cost. However, the purchasing cost is a constant as long as purchasing price is not a function of order quantity.
5. The total cost function is not sensitive to the order quantity near the EOQ. See Figure 7.9.
6. At the optimum, the total setup cost is the same as total holding cost.
7. The annual purchasing cost Dc is constant. It is typically much higher than the setup cost plus the holding cost. This is not true if cost is a function of Q, such as in quantity discount.
8. Buyers of often face challenges that EOQ < minimum order quantity (MOQ). MOQ is often vendor's EOQ, or some convenient quantity based on historical reasons. You should work with vendors or expand your market to increase your own EOQ.
9. In reality, the strict assumptions of EOQ rarely holds. However, EOQ is quite robust, meaning it provides excellent guidelines on order quantities even when some assumptions are violated.

Lean has many dimensions. One of them is directly related to EOQ. If EOQ is high, the system will carry a lot of inventories, work in process or finished goods, and by definition is not lean. What can you do to reduce EOQ?

You can reduce *K*, *D*, or increase *h*. No one in their right mind would try to actually decrease demand or increase holding cost, although I have encountered students who stated such solutions in test! However, there is something to consider about increasing the accounting of the holding cost. It is not about increasing the actual holding cost but about how to account holding cost accurately. Many companies use the discount rate, internal rate, return, or other financial measures to account for holding cost. This can be much lower than what the holding cost should be. HP published a case study on inventory driven-cost on electronics. It found that the traditional way of accounting holding cost only captured a small portion of the total inventory-driven costs (Gallioni, Montgros, Slagmulder, Wassenhove, and Wright, 2005). The other cost can be obsolescence, rebate, returns, price drop, handling, storage, shrinkage, and many others.

The most effective way to reduce EOQ drastically is through the reduction of setup cost. Toyota started just-in-time (JIT) production. JIT demands small batch sizes. The first thing Toyota did to promote JIT was setup reduction through an effort called Single Minute Exchange of Dies (SMED). It was related to the change of dies for metal stamping of car body panels. The panels are shaped with very large and heavy die sets. It would take several hours to change and adjust a new die set. The stamping process, on the other hand, is very fast, at a few seconds a piece. After the die change, one would naturally stamp many panels, which leads to a lot of inventory and WIP. The SMED effort led to tremendous reduction in setup time and the batch sizes.

5.2 Lot-size scale effect from the demand side

Supply-chain texts often focus on the inventory policies with given demand. A different question is: what is the inventory cost if the demand increases? We know that the inventory cost per unit time increases with the demand. I have heard consultants apologize to the client that the increase in demand will increase inventory cost, which can be shown in the cost function. However, is it a good measure? Would inventory cost per unit per unit time be more meaningful?

You can verify that the total inventory-related cost at EOQ by substituting the optimum order quantity in the cost equation, or

$$C(EOQ) = \sqrt{2KDh}\ \$/\text{per unit time}$$

To convert this quantity to per unit per unit time, we can divide this by the demand per unit time *D*, the average inventory cost per unit per unit time is

$$AC(EOQ) = \frac{\sqrt{2KDh}}{D} = \sqrt{\frac{2Kh}{D}}\ \$/\text{unit/unit time}$$

This is a decreasing function of the demand! Therefore, increasing demand not only increase the revenue, but it also reduces the average inventory cost per unit at the same time.

Example

Due to population growth, the streamlined process, and reputation, the auto repair shop quadrupled their demand for the filter to 800 per week. The ordering cost stays at $100 per order, the cost of a filter is still $16 with the same holding cost at $0.04/week. What is the approximate inventory-related cost per unit per year before and after near EOQ? Assume 52 weeks in a year.

Solutions

The inventory cost per unit per year before is

$$C_0(EOQ) = \sqrt{\frac{2*100*0.04}{200}} * 52 = \$10.4/\text{filter/year}$$

At the higher demand level

$$C_1(EOQ) = \sqrt{\frac{2*100*0.04}{800}} * 52 = \$5.2/\text{filter/year}$$

This means that quadruple the demand can reduce the inventory-related cost per unit per unit time by a factor of two!

5.3 Quantity discount

It is common, especially in the United States, for firms to offer a quantity discount. There are many reasons that sellers would provide a quantity discount. It could be that it passes its cost savings to you. For example, most sellers will charge you much less if you buy by unit load such as a case, a pallet, or a truckload. The quantities less than those unit loads need more handling in order picking, packaging, or transportation per unit. A forklift can handle thousands of items on a pallet in one trip. TL cost less than LTL. Another reason can be to compete for market share. Customers that buy more of their products would buy less from others. It can also be used to smooth the demand to manage its inventories and production.

With quantity discount, the purchasing cost must be included in the inventory-related cost in the model

$$min\, C\,(Q)/unit\ time\ = K\frac{D}{Q} + h\frac{Q}{2} + Dc(Q)$$

Where $c(Q)$ is a decreasing function of Q. In the EOQ characteristics, we observed that first 2 terms in this equation are often much lower than the third term near the saddle point Therefore, the reduced price can have significant impact on the total cost.

In a three-level all-unit quantity discount, the purchasing cost for pieces, caseloads, and pallet loads can be expressed as

$$c(Q)\ \$/\text{unit} = \begin{cases} c_0, & 1 \le Q < Q_1 \\ c_1, & Q_1 \le Q < Q_2 \\ c_2, & Q_2 \le Q \end{cases}$$

Here, cD is no longer a constant. You must include it in the inventory cost per unit time, or

$$C(Q)\ \$/\text{unit time} = \begin{cases} K_0\dfrac{D}{Q} + h\dfrac{Q}{2} + c_0 D, & 1 \le Q < Q_1 \\ K_1\dfrac{D}{Q} + h\dfrac{Q}{2} + c_1 D, & Q_1 \le Q < Q_2 \\ K_2\dfrac{D}{Q} + h\dfrac{Q}{2} + c_2 D, & Q_2 \le Q \end{cases}$$

The three expressions can be represented by three-piece-wise convex curves—red, black, and green, respectively—shown in Figure 7.10. The red curve is valid below $Q < Q_1$, indicated by the red arrow. The black curve is valid $[Q_1, Q_2)$, and the green when $Q \ge Q_2$. The minimum cost solution will be the minimum among the valid EOQ(s) and or/a boundary location. In the situation shown in Figure 7.10, $Q^* = Q_2$ because the cost is lower than the valid EOQ on the black curve. The EOQs on the red and green curves are invalid.

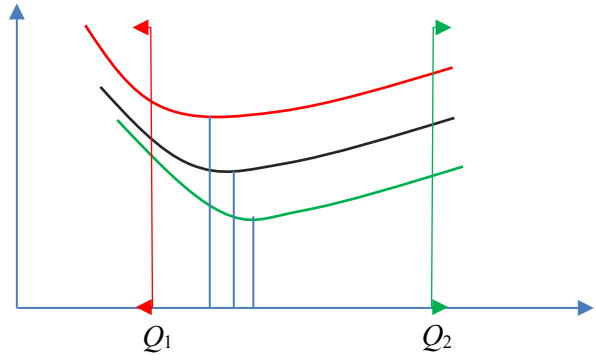


Figure 7.10. Different cost curves and realizable regions at different costs for different quantities.

For more information on quantity discount, please refer to other references.

6. ECONOMIES OF SCALE IN INFORMATION TECHNOLOGY

The software industry has a high development cost but very low distribution cost per license. This includes enterprise software, personal software, and games. Consider Microsoft, Google, Oracle, SAP, Manhattan Associates, Nintendo, and Sony. The social media industry also enjoys significant economies of scale and also a strong network effect. Consider Facebook. You want your friend to be in Facebook, not other platforms. In such an industry, winner take it all is prevalent.

7. ECONOMIES OF SCALE IN BARILLA CASE

Barilla takes full advantage of the economies of scale in many ways. It adopts high throughput and low marginal cost production systems to gain competitive advantage. They also batch the production to minimize the setup and holding costs per unit. Due to high setup cost, they used the quantity discount to promote the products planned for production. These strategies have helped Barilla to grow into a dominant supplier. However, there have been issues that need to be addressed. More discussion to follow in the following chapters.

8. ECONOMIES OF SCALE AND SUPPLY OF HUMAN NEEDS

One of the major motivations to grow, to merge, or to acquire is to enjoy more economies of scale. Even large supply-chain firms such as smart phone makers, transportation service providers, and brick and mortar or e-commerce retailers still want to grow to enjoy more scale effect.

8.1 The benefits

Online retailers such as Amazon are big and are getting bigger. They changed the lives for many. The larger scale allowed the retailers to adopt computerization and automation to reduce the marginal cost. The lower cost allows them to lower their prices, and benefits the consumer with tremendous choices, convenience, and fast delivery. It also created indirect and direct competition with traditional retailers that incentivizes the retailers to reduce the cost further. Walmart's 2020 Q1 e-commerce sales were $8.3 billion, or 6.2% of the net sales, compared to $4.7 billion the same period of the prior year (Walmart 10-Q report). The scale, and the lower prices, allow the consumers to use their income for additional needs, such as healthcare. This is especially important for the low-income population.

The consolidation with new technology may lead to higher quality, lower emissions, lower waste, and lower environmental footprint. The lower production cost allows the firms to invest to improve worker's work environment to improve their health. All these improve supply to human needs.

8.2 The drawbacks

The dramatic cost reduction measured by amount of time it would take to do something leads to oversupply of certain goods, such as fashion, food, and electronics. You can observe your consumption of your favorite snack. If you buy a small package, what is your consumption rate? If you buy a large package at lower cost per unit, what is your consumption rate? Getting larger packages at lower cost can also lead to waste.

The low costs by the existing large suppliers make it difficult for the smaller players to enter the market, even with innovative ideas. Take food, for example. We know that organic food, fresh food, and pastured chicken are better for human health needs. In some cases, they are more sustainable. However, traditional large-scale farming with fertilizers, herbicides, and genetically modified seeds enjoys tremendous advantage in economies of scale in production, transportation, storage, and retail over organic food, fresh foods, and pastured chicken.

The economies of scale also put the new entrants at disadvantage even if the new entrant provides better product or service for human needs. As a different car supplier, Tesla went to scale quickly to avoid the competitive disadvantage for their popular models 3 and Y. Their plants start off large to enjoy the scale. Even though they lost the tax credit because they have sold more than 200,000 cars, its prices are still competitive with brands that still enjoy the tax credits. However, few new entrants can achieve this. Scale effect is strong. Without the scale, the cost is often too high.

8. ECONOMIES OF SHARING AND ECONOMIES OF SCOPE

Variety is the spice of life. With the development of modern technology, customers demand more variety, customization, or even personalization. The customer preferences in products and services vary among many dimensions such as size, color, functionality, durability, price, quality, style, taste, health implications, social impact, and environmental footprint.

However, adding variety in supply increases supply-chain complexity. It has a potential to increase revenue but also increase the cost. Firms must develop strategies to limit the cost increase and maximize the revenue increase when adding variety. Sharing is at the center of these strategies.

One strategy is to optimize the number of assortments or variety of products or services. The less variety enjoys more economies of scale and simplicity in the supply chain. However, this can limit the market and revenue. There is also choice overload when too many choices are offered. There exist optimum levels of variety that maximize the profit, considering both the cost and revenue.

The second strategy is to optimize the supply chain with minimum variety of resources shared by most products or services. In this strategy, different products or services must share the limited set of resources. This simplifies the management, planning, and maintenance of the resources. However, the resources may not be the best choice for some of the products, such as too big or too small. It may also limit the ability to supply certain types of demands.

The third strategy is to minimize the variety of input materials or customers shared by most variety of products or services.

The fourth strategy is to design a flexible supply chain that is capable to supply a large variety of different products or services to enjoy the economies of scope. The computer, robotics, and artificial intelligence have provided unprecedented support for the flexible supply-chain strategy.

There is also a strategy based on the sharing economy to maximize the utilization of the underutilized resources to satisfy demand that cannot be economically satisfied with a non-sharing economy.

1. Optimize the Product Assortments Shared by Variety of Demand

Clients and customers often demand something different from what a supplier offers in many dimensions, such as size, color, functionality, durability, price, quality, style, taste, health implications, social impact, and environmental footprint. Internally, the marketing would always ask supply-chain professional to add more variety to increase revenue, such as in Barilla's case.

In *My Life and Work*, Henry Ford stated:

The salesmen, before I had announced the policy, were spurred by the great sales to think that even greater sales might be had if only we had more models. It is strange how, just as soon as an article becomes successful, somebody starts to think that it would be more successful if only it

were different. There is a tendency to keep monkeying with styles and to spoil a good thing by changing it. The salesmen were insistent on increasing the line. They listened to the 5 percent, the special customers who could say what they wanted, and forgot all about the 95 percent, who just bought without making any fuss. No business can improve unless it pays the closest possible attention to complaints and suggestions….

1.1 Cost, revenue, and profit with respect to variety

Consider a food truck. The supply variety, or assortment, is shown in the menu. The owner must have already put in a lot of thought into the design and selection of the menu items, ingredients, truck, equipment in the truck, and personnel needed to run the operation. There will always be customers who want something different. The owner would have to decide to what to supply.

To increase in variety, the firm may need to add resources and modify its supply chain and additional input materials; both lead to more cost. The firm would first add the varieties that are similar to the existing products. Such variety enjoys some sharing of the existing supply chain. Further additions will become less and less similar to the existing products and lead to a higher rate of increase in costs. In the food truck example, the owner may first add an item that shares more in common with the existing menu items, such as a different mix of current ingredients. However, adding an item that requires different ingredients or different cookware may lead to higher cost. Therefore, the cost as a function of variety should be increasing and convex, shown in the long dashed red line in Figure 8.1.

More variety should create more market and revenue. A firm can first add the variety with the least increase in cost and most increase in revenue. However, the continued increase in variety will lead to diminishing return. In addition, the added variety may also cannibalize the market of the current varieties, unless the addition can enhance the sale of current variety. Therefore, revenue as a function of variety should be an increasing and concave function. The profit is the difference between the revenue and cost, also concave, with a peak, illustrated as a solid green line in Figure 8.1.

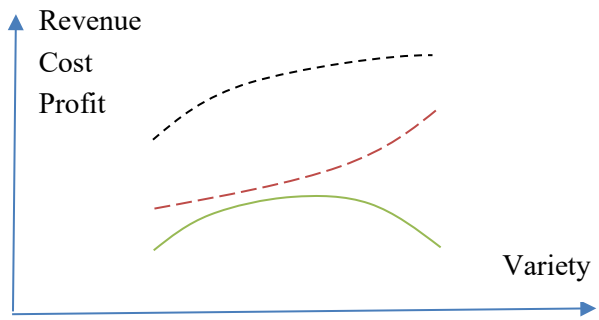


Figure 8.1 Illustration of revenue, cost, and profit in sharing assortment.

1.2 Sharing in a limited product assortment

Consider Google's Pixel phones. There is a small number of assortments in functionality, size, and color. As a result, potential customers may not get their exact preference. For example, Pixel 4 only comes with three models: 4, 4a, and 4XL. The most popular model, 4a, only comes in a single color, black. Maybe 30% of the customers prefers black. An additional 50% of people may sacrifice their color preference to their preference for memory, speed, size, and price. The rest of the potential customers would just get another brand. If Google reached the point that the average cost of the single color started to go flat, it could always add a second color. Adding a second color would add cost in the production system and logistics system in terms of setup, handling, transportation, and inventories. They must also consider if the new size cannibalizes the demand for the current size. Other dimensions easier to associate can be weight and speed.

The marketing department is incentivized to increase total quantity and revenue. One way to increase quantity is to add more assortments.

The marketing department does not have to deal with the additional cost and complication in the supply chain. It is up to the supply-chain professional to show these issues, preferably quantitatively.

1.3 Choice overload

Even the premise that more SKUs increases revenue may not always be true. Too many choices can also slow down the selection process, lead to regret with the selection, indecision. and even no sales.

The book *The Paradox of Choice*, by phycologist Barry Schwartz, argues that too many choices add anxiety to the consumers, slow down the selection process, and can even lead to making no selection at all. Even when the customer made a choice among many assortments, they become less satisfied due to second-guessing on the close choices not chosen.

Shernev, Bockenholt, and Goodman performed a meta-study to analyze 99 observations with a sample size of over 7,000 in prior studies. They identified four factors that create choice overload when a customer selects among assortments. These are:

1. Choice set complexity. The complexity in the choice decision grows with the number of choices. The choices can be along different dimensions such as color, size, style, health implications, environmental impact, durability, functionality, and price. Each product or service is the realization of certain combinations along these dimensions. The more choices, the more complex is the set of combinations.
2. Decision task difficulty. The complex set of choices with complex combinations in different dimensions make the decision difficult.
3. Preference uncertainty. The preferences are based on the knowledge among all dimensions and the understanding of self-preference. However, most buyers have limited knowledge among many dimensions of many choices. This leads to uncertainty, decision deferrals, or weaker preferences for the choice made.
4. Decision goal. To minimize the effort needed to make a choice. More choices will most likely increase the effort needed.

The authors also found that four factors or measures from prior studies are equally powerful: satisfaction/confidence, regret, choice deferrals, and switching likelihood.

1.4 The inventory-related cost of assortment

Consider inventory-related costs associated with variety. Some firms adopt a strategy with limited varieties while others with many varieties. The big box retailers, such as Walmart or Target, operate many super-centers. They are large stores with many product categories—tens or even over one-hundred-thousand categories or SKUs.

Aldi, one of the budget German-owned chains with limited assortment, adopts a very different strategy. They run small stores, with only popular categories of products, and with a limited assortment within each category.

Costco also offers much less variety than other big box stores. They are also very successful.

In 1997, Walmart expanded into Germany. By 2006, they ran 85 stores. However, their sales were lower than expected. They initially thought that maybe they did not know German consumers. To their surprise, they found their prices were higher than some of their German competitors such as Aldi. How could that be? We can compare the inventory-related costs to understand how a small store such as Aldi can beat Walmart on price.

Example 2

The annual total sales of toothpaste in a particular Aldi store are the same as a particular big box store of toothpaste by weight W. Aldi has 3 SKUs while the big box store has 48 SKUs. Please provide rough estimates for:

1. Inventory cost related to cycle stock (we will discuss safety stock later).
2. Space requirement for cycle stock.
3. Impact of economies of scale.
4. Other factors.

Discuss the drawbacks of limited variety.

Solutions

Assumptions

1. All EOQ assumptions.
2. The total demand represented by weight sold is W for both the big box store and Aldi.
    a. Let the weights of the three Aldi SKU demand be W_A, W_B, W_C.
    b. Let $w_i, i = 1 - 48$ be the weight of each of the SKUs from the big box store.
3. The setup cost/restock of one SKU is the same, regardless of quantity or firm.
4. The holding cost per unit weight per unit time is the same across SKUs and firms.
5. Both stores' cycle stocks are near their EOQs.

1. Inventory-related cost. From the information given and assumptions, we can link the demands in the big box store and Aldi with

$$W = \sum_{i=1}^{48} w_i = W_A + W_B + W_c$$

Consider SKU_A in Aldi's "high-velocity" item. We can find 16 high-velocity SKUs from the big box store SKUs that satisfy this condition

$$\sum_{i=1}^{16} w_i \approx W_A$$

Since the inventory cost is insensitive when the order quantity is near EOQ, the total costs can be approximated by

$$\sum_{i=1}^{16} C(EOQ_i) = \sum_{i=1}^{16} \sqrt{2Kw_ih}$$

Let *w*' be the average demand in weight across 16 SKUs. Then the sum of weight is equal to 16 times the average weight.

$$\sum_{i=1}^{16} w_i = 16w'$$

At EOQ, the sum of setup cost and holding cost is equal to $\sqrt{2KDh}$. Therefore, for the big box store, we have

$$\sum_{i=1}^{16} C(EOQ_i) = \sum_{i=1}^{16} \sqrt{2Kw_ih} \approx 16\sqrt{2Kw'h}$$

When the weights of 16 SKUs are identical, the two quantities are the same. When the weights of different SKUs are different, the term on the left-hand side is lower. Since we select the weights of selected SKUs are similar, the two sides should be similar. In Aldi, there is only one SKU_A with demand equal to all 16 SKUs in the big box store. We have

$$C(EOQ_A) = \sqrt{2KW_Ah} \approx \sqrt{2K(16w')h} = 4\sqrt{2Kw'h} \approx \frac{1}{4}\sum_{i=1}^{16} C(EOQ_i)$$

This is only one-fourth of the inventory cost on cycle stock in the big box store! Since the weights in the big box store are not identical, the savings may be lower but still very significant. You can use the same rationale for W_B and W_C.

2. Space requirement. For the cycle stock, you can do the same to find the relationships between total average cycle stock.

$$\sum_{i=1}^{16} EOQ_i = \sum_{i=1}^{16} \sqrt{\frac{2Kw_i}{h}}$$

If the weight across the SKUs is the same as w', we have (this is a strong assumption, we will explain later),

$$\sum_{i=1}^{16}\sqrt{\frac{2Kw_i}{h}} \approx 16\sqrt{\frac{2Kw'}{h}}$$

Therefore,

$$EOQ_A = \sqrt{\frac{2KW_A}{h}} = \sqrt{\frac{2K(16w')}{h}} = 4\sqrt{\frac{2Kw'}{h}} = \frac{1}{4}\sum_{i=1}^{16}\sqrt{\frac{2Kw_i}{h}}$$

This means, the cycle stock in Aldi is only one-fourth that in the big box store. This will lead to space savings.

3. Economies of scale

In handling in logistics throughout the supply chain, the higher volume per SKU will mean more pieces in the handling unit: from mixed SKU cases to single SKU units, caseloads to mixed pallet loads, mixed pallet loads to single SKU pallet loads, and from LTL to TL.

In procurement, Aldi enjoys approximately 16 times larger volume per SKUs. This can lead to bigger quantity discount from the supplier.

4. Other factors

Since Aldi's minimum handling loads are mostly caseloads, they can actually use the cases as their supplemental storage unit on their shelves. Therefore, Aldi design their cases in a way that the can be opened for access directly from the shelf. This saves the task of taking out the items from the shipping cases at the store and reduces the cost in shelves. Aldi also has fewer SKUs to update, manage, and so on.

5. Drawbacks of fewer SKUs

There is a potential that some customers really want to get the SKU Aldi does not have. Some of them may be regular Aldi shoppers but go to other store for the favorite toothpaste, probably at higher cost. Some of them may not come to Aldi at all. As a supply-chain professional, you must judge the pros and cons in your business model.

Therefore, the products or services assortment is an optimization problem that involves cost, revenue, market share, customer psychology, and sales speed. A supply-chain engineer should consider all these factors in developing the strategies.

2. SHARING RESOURCES

There are many types of resources, including time, space, equipment, facilities, and human resources.

2.1 Sharing over time and in space

For most products and services, demand for one product may vary over a year, month, week, or even in a day. For example, take the products for winter sports and summer sports. Many firms purchase consumables at certain time in the month to balance their budget. Demand varies in a week due to weekly work, pleasure, exercise, and religious activities. Demand varies in a day to fit in people's daily routine. In the food truck example, the owner can design the system that can work in winter and summer, different days in a week and breakfast, lunch, and dinner. The owner can also design the system so that it can used for other purposes on the weekend.

In brick-and-mortar retailing, the cost of well-designed space in prime locations can be very high. It would be a great waste to have the store idle for certain parts of the year, month, week, or day. The supply chain professional should work hard to find complementary products or services to share the expensive space. In a sports store, the gear for winter sports and summer sports can share the same space. Costco has the highest sales per unit time and square foot among most merchandizers. It carries the popular staple products that are available year long. It also allocates some space for seasonal items or flash sales. The window of availability is short; the customers know that if they do not get the product they like at once, the product may not be available on the next visit. This "artificial scarcity" strategy increases their sales rates! More about this in the chapter uncertainty.

In services, a firm can simplify its operation times so that the customers would share more resources. For example, the operating hours of banks are often limited. The customers would adjust their own schedule to visit during the operating hours. This strategy increases the resource utilization. However, such strategy decreases customer service. Some banks would operate more hours to compete. The drawbacks can be the loss of customers, lost capacity in idle times, and so on. The sharing of time and space as a function of variety also follows the trend of cost, revenue, and profit shown in Figure 8.1.

2.2 Sharing of equipment or facilities

The production and or service often needs equipment: machines, printers, X-ray machine, cash register, trucks or busses. For each function, there are many types with different functionality, throughput capacity, ease of operation, and cost. A firm can adopt many. Each is most suitable for each type of demand, or a single one. This is similar to the product portfolio management but with equipment.

Let us consider some industries a typical consumer can relate to. Consider school buses. These are unit load services. Some routes need more seating capacity than others. One strategy is to unify the equipment of a single size that is suitable for most routes. This simplifies management, maintenance, and many factors. For the routes with lower demand, there will be waste of capacity. For routes where one bus is insufficient, a second bus can be used. Note that the school bus is mostly used in the morning and afternoon. The operators often work hard to get additional demand to share the available capacity. Public transportation shares the similar situation. It is more challenging because the capacity required in rush hour is much higher than during other hours. They often schedule more frequent services during rush hour. Transportation companies have similar issues. Classrooms and trucks come in different sizes. A firm must determine what strategy to adopt. Below, we look into the details of airlines.

Delta Airlines is one of the largest airlines in the world, which provides many classes of services to many international and domestic destinations through its hub-and-spoke network. In its network, there is a lot of demand between hub cities that requires large planes to fly multiple times a day. In some of its spokes, such as from Atlanta to Brunswick, Georgia, it only needs to fly a small plane with limited scheduled flights. To provide such services with the right capacity and right class, Delta operates over 900 planes including Airbus 220s, 319s, 320s, 321s, 330s, 350, and Boeing 717, 737s, 757s, 767s, 777s, and MD 88 and 90 (it retired its Lockheed 1011 in 2001 and 747 in 2017). The wide range of equipment allows Delta to utilize the equipment with right seating capacity, cruise range, and classes of services to fit the right routes for all types of customers. Delta had moved up among the major airlines in service, revenue, and many other measures before the Covid-19 pandemic. However, this strategy leads to higher cost for training, maintenance, and spare parts. This strategy also adds complexity in scheduling, training, staffing and support.

Southwest Airlines is a budget airline that provides mostly domestic services and some resort locations in Mexico and the Caribbean. It is the third-largest airline in the United States, with about 750 airplanes in 2019. They are all Boeing 737s! These planes vary only in model year and length. They share most of the other components. This reduced the cost for the training, maintenance, and spare parts. This strategy also reduces the complexity in scheduling, training, staffing, and support. However, these planes are sometimes too big, other times too small, or do not provide the desired class of services to certain customers. The equipment sometimes is too big while other times too small.

2.3 Sharing in unit load in material handling

Material handling is defined as the movement, protection, storage, and control of materials and products throughout manufacturing, warehousing, distribution consumption, and disposal by MHI. To accomplish these, material handling is often done in unit loads such as shipping cartons, cases, totes, pallets, skids, or containers and trucks. In order to reduce the number of trips and reduce the space requirements, different products can share the same unit load, such as mixed cases, mixed pallets, and less then truckload (LTL) shipments. The LTL services use cross-docks to maximize the capacity utilization of truck unit loads. A service closer to consumers is the merge in transit, when the delivery person merges customer orders from different sources into a single delivery. This reduces the cost of delivery and adds convenience for the consumer.

2.4 Sharing of personnel and cross-training

Firms may require different types of personnel for different products or services. One example is service at a cafeteria. The cafeteria opens from 11 to 2. The peak time is from before noon to 1. Before the peak time, a lot of personnel are needed in the kitchen to get the food ready. At the peak time, a lot of personnel are needed as cashiers. After the peak time, more people are needed for cleaning, clearing, and other services. Instead of staffing for the peak demand for each category of tasks, one strategy is to share the personnel among different tasks.

Example

A third-party logistics (3PL) company runs a distribution center that provide services to one major client. The annual cost due to investment of the building is $2 million per year. The annual

investment cost due to investment in storage, handling equipment, and trucks is $2 million. The cost of labor is $1.5 million per year. It found a second client with similar services needs and similar revenue. They can build another distribution center with the same structure or modify the existing system to accommodate the new client. The cost to invest in the existing warehouse management system (WMS), storage and handling equipment, and trucks is $2.5 million per year. The added labor cost will be $1.0 million. Please compare the two systems.

1. What are the total costs of each system? What would you recommend?
2. What are the other factors to consider?

Solutions

Assume the information provided is sufficient for this rough-cut analysis.

1. If the 3PL would set up a second distribution center to handle the second client, the cost would double, or

$$TC = 2 * (2 + 2 + 1.5) = 11$$

If the 3PL would modify the existing facility to let the two clients share the services, the total cost will be

$$TC = (2 + 2 + 1.5) + (2.5 + 1.0) = 5.5 + 3.5 = 9$$

The economies of scope, or the savings from sharing, is 11 – 9 = 2 million. The savings come from the building share, ($2 million) and labor share ($0.5 million). However, the new storage and material handling system of 2.5 million is $500,000 higher than the first option.

2. There are other factors to consider. If the 3PL has potential to serve more clients, the single location may have maxed out. The new DC may be in better location for other services.

3. SHARING IN INPUT AND WORK IN PROCESS

A firm in the supply chain converts inputs into outputs as the products or services.

3.1 Select inputs that can be shared by most products

The more the products or services can share the same input, the less input a firm must deal with, such as screws, sheets, chips, glass, and wires. The benefits are in the management, inventory, transportation, storage, space requirements, and many other facets in the supply chain. This strategy can be related to product design, and other factors already discussed earlier.

3.2 Postponement in differentiation

The above is true for work in process (WIP) within the firm, since different products and services may share the same WIP. The processes that cut, drill, shape, weld, and glue will customize the raw material toward a smaller set of products, and are no longer useful for other products. The demand can be unpredictable. This can generate waste. Therefore, in supply-chain design, the postponement of differentiation is very important. HP started to coin the term postponement of

product differentiation. This allows the material to stay in general form as close to the time of actual demand as possible to maximize the sharing.

For this reason, more and more light processes are moved downstream closer and closer to the end-user. Distribution centers provide more and more light duty value-added functions, such as labeling, light assembly, or putting customized accessories into customer order containers. For example, Whirlpool designed refrigerator doors that can be mounted to open either to the left or right. The door can be installed at the distribution center before delivery, or even when being installed at homes. Even the store can provide some services such as the length of pants, initials on shirts, and so on.

3.3 Modular design

The products and services can be designed as modules. The assembly of different modules can form different products with different functionalities. Many different computers, cell phones, chairs, and types of software, with different capacity, size, and functionality, can share the same modules.

Example

A firm supplies a tool. It offers a single yellow color, based on market studies. The total revenue for the tool is $1 million per year. The margin on the product is 10%. Later, a competitor starts to supply a tool with similar function in blue. The firm's revenue decreased. The marketing department in the firm suggests adding three more colors to maintain the market and revenue. The total cost of the raw materials in yellow was $30,000 per year. The total cost for all colors at a smaller scale will be 2% higher due to quantity discount, and 10% higher in transportation, inventory, and management. The additional tooling cost will be $3,000 per year. The additional changeover cost will be $8,000 per year. The added complexity will require software changes, the manager's attention, and worker training. The total is estimated at $10,000 per year. If new colors are not added, the revenue on yellow only will drop 70%.

Questions

1. What will be the total added cost due to the addition of the new colors?
2. How would you present your analysis to management?

Solution

This is a case of with cannibalization. Assume that the problem statement has captured all the major demand and cost information. The annual equivalency costs are sufficiently accurate for the purpose of this rough analysis.

1. The extra cost items per year

The total added cost due to materials is 30,000 * (0.02 + 0.10) = 3,600.

The total extra cost in tooling is 3,000.

The total added changeover cost is 8,000.

The total cost due to software, management, and training cost is 10,000.

The total extra cost to add the three colors, or the sum of the above, is 24,600 per year. This represents 2.46% of the revenue but 24.6% of the margin.

2. If the colors are not added, the revenue will be drop by 30%. If the margin remains at 10%, the total margin will be $100,000. A 30% drop will be $30,000, about 20% higher. Therefore, the addition of three colors is slightly better based on these numbers alone. In practice, there are many other considerations; some of these will be discussed in the following chapters.

4. ECONOMIES OF SCOPE

Adopting the product-sharing and resource-sharing strategies, a firm will always have the potential to lose customers.

The economies of scope refer to the fact that the average cost to supply another variety lowers the average cost per unit among different products. The cost reduction is because the flexible supply chain can produce different products as if they are similar or even the same. The computer, robotics, and artificial intelligence have provided unprecedented support to economies of scope through flexibility. For example, 3-D printing provided the ultimate flexibility in generating almost any types of geometry required by the products or services. The limitations today are in the materials and sizes. However, limitations are being broken all the time.

4.1 Modeling of economies of scope in supply chain

In order to produce product or provide service, a firm must invest in resources such as getting space and purchasing equipment to product products or provide the service. The actual costs are always very complicated. However, we can use a few important parameters to get important insights into economies of scope. Let,

n *The number of different products to supply.*

I^S *Shared investment cost among all types of products.*

I_i *Dedicated cost to produce product* $i, i = 1, 2, \dots n$.

c_i *Total incremental cost per unit for product type i.*

q_i *The quantity of product i produced.*

AC *The aggregate average cost among all product varieties.*

Using these parameters, we can estimate the average cost of each product among all product types to be

$$AC = \frac{1}{\sum_{i=1,\dots n} q_i}\left\{I^S + \sum_{i=1,\dots n}[I_i + c_i q_i]\right\}$$

The denominator is the total number of the products produced among all product types. The first term in the bracket is the total investment cost shared by all product types. The second term is the summation of the total costs for each product not shared. For each product, there is an investment-related cost, which can include facilities, machines, personnel, and so on. The second term in the summation is the production cost per unit times the production quantity for each product type.

Consider firm A that produces two products in two separate systems versus firm B that produces the same two products with a flexible system. Firm A does not have shared facility, therefore I^S

= 0. The total cost is the summation of the cost using two dedicated system. Firm B will incur investment cost I^{SB}. Typically, $I^S >> I_i$. The average costs of these two products depends on the relationships among I^S, I_is and c_is. In Figure 8.2, the dashed lines are for firm A and the solid lines are for firm B.

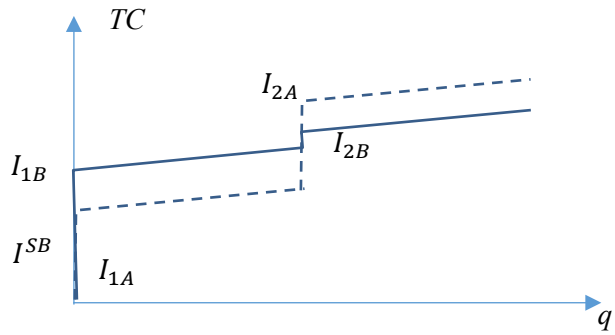


Figure 8.2. Comparison of dedicated systems versus flexible or shared systems.

Therefore, the revenue, cost, and profit curves of flexible systems versus variety will have the trend shown in Figure 8.3. As the variety increases, the cost increase is slow due to flexibility. The revenue and profit increase will be increasing within the range of the system scope, as shown in Figure 8.3. This is different from the sharing in assortment, resources and input because the flexible system is designed to be shared in a wide range of products and services.

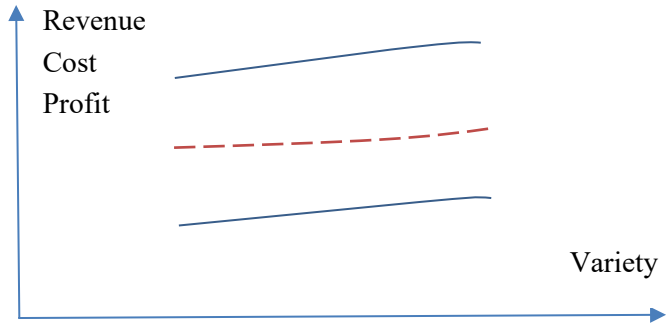


Figure 8.3 Illustration of revenue, cost, and profit as a function of variety with flexible systems.

Example

A firm has been making two types of shells for recreational boats with two dedicated production lines. The firm wants to add five other types of shells to its product lines. The demand for the five new types is estimated to be up to 500 per year. The investment cost for the dedicated line is

20,000 per line per year. Its capacity is 1000 per line. The marginal cost for making one shell is 50 per shell. Nowadays, there are flexible systems that can be used to make any types of shells. One flexible system cost is 90,000 per year per line, with a capacity of 3000 per year. The marginal cost is 48. The changeover cost between different types per year is 1000.

1. What are the total and aggregated unit costs adding the five new types?
2. Please discuss your solutions.

Solutions

Assumptions

1. The information given is sufficient for approximation purposes.
2. The aggregation of different shells will not result in significant errors.
3. The two current lines will continue to be in service, and will not be considered.

1. The aggregate total number of new shells to be produced, regardless of its type is

$$\sum_{i=1,\dots n} q_i = 5 * 500 = 2500$$

The total annual cost with five added dedicated lines will be

$$TC^D = I^S + \sum_{i=1,\dots n}[I_i + c_i q_i] = 0 + 5 * (20000 + 50 * 500) = 225{,}000$$

We replaced summation with multiplication because they are identical. The annual cost with five added dedicated lines will be

$$AC^D = \frac{225{,}000}{2500} = 90$$

The added demand is equal to 4000 for the right new shells. Since the flexible line capacity is 2000, the firm needs to add two flexible lines. The total annual cost of the existing system with two added flexible lines will be

$$TC^F = I^S + \sum_{i=1,\dots n}[I_i + c_i q_i] = 90000 + 5(1000 + 48 * 500) = 215{,}000$$

$$AC^F = \frac{215{,}000}{2500} = 86$$

The flexible system costs 4.4% below the dedicated systems.

2. In practice, there are many other considerations. We used this simple model to perform a rough-cut approximation to illustrate how to assess the cost or benefit of flexibility. The flexible system often cost more in initial investment. However, it reduces costs in changes

between varieties. In reality, there are other factors to consider: economic life of the lines, net present value; and what to do with the two current dedicated lines. The dedicated line has more extra production capacity. What if the demands are higher than forecast? The flexible system is flexible, and it can be used to produce new designs not envisioned today.

4.2 Computers and humans support flexibility

Human beings are very flexible. A handy man or woman can do many different things. An electrician, a plumber, or a gardener can do wonders with a small set of tools. A garage can fix a large variety of problems of many brands and models within limited space and some generic tools. Even in Amazon's most automated generation order-fulfillment centers, the last step to take an item from the storage container and place it in a customer order container is still mostly manual. We can produce self-driving cars, but we are still challenged to produce a cost-effective robot that is able to pick up a marker, a pair of socks, a hat, a water bottle, a stuffed animal, nail-clippers, and a book. The sharing of generic tools and human capabilities among many useful capabilities provided the flexibility in manual systems.

Computers are very flexible. Any computing device is some combination of CPUs, memories, and communication components. Coupled with sensors and actuators, they can be programmed to perform many types of functions in the supply chain: monitoring, moving, control, plan, execute, or interface. The sharing of the resources of many useful capabilities provided the flexibility in computer systems.

Artificial intelligence, robotics, and 3-D printing offer the ultimate in flexibility and adaptability.

4.3 Characteristics of different supply systems

Consider manufacturing systems. The term "job shop" is used refer to highly manual systems that can process many different types of jobs with some general-purpose tools such as drills, cutting machines, grinders, and some manual tools. It hires skilled personnel who can perform many different tasks. In such systems, most of the investments are to be shared by many different jobs. The setup time for each job can be long and setup cost for each job can be high. The incremental cost per job tends to be very high because it requires highly skilled worker for long period of time. Imagine having a cooling system or oil leak in a car. A mechanic must first determine the problem, order the parts, set up the car and tools, and work on it for a long time. The cost of labor is often much higher than the parts.

On the other extreme is the transfer line or dedicated flow line. This can be for the production of cereals, carpet, or glass. The system must have a very high throughput rate, and high quality through multiple stages, with rigid requirements on cleanliness, temperature, humidity, and so on. Some of these lines are often dedicated to a single product. The total cost per unit is low.

These two systems are depicted in the upper left and lower right in a variety n versus volume q_i chart shown in Figure 8.4.

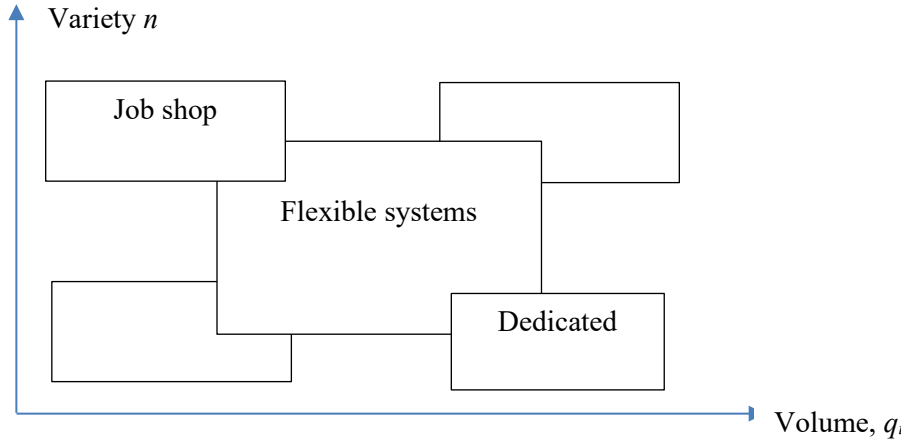


Figure 8.4. Job shop, flexible and dedicated systems on variety versus volume chart.

The flexible systems add flexibility to the dedicated systems to allow different jobs to share the same system with high speed and lower cost. In the advanced economies, the cost of computing and robotics are lower than in the developing economies. Their customers may need more customization and even personalization, such as putting initials on a shirt or a bag. In addition, their workforce is more familiar with computers and robotics. As a result, flexible systems are most suitable in the developed economies.

Table 8.1 shows the cost characteristics of different types of systems. Where H, M, and L are used to indicate high, medium, and low. The entry "*" for AI, 3-D and robotics indicates the cost can vary widely. In the design of the supply systems, a firm can consider these characteristics in formulating their design strategies.

Table 8.1 The cost characteristics of different types of systems.

System \ Performance	Throughput	AC	Shared investment I^S	Dedicated cost I_i	Marginal cost
Job shop	L	H	M	M	H
Dedicated	H	L	L	H	L
Flexible systems	M	M	H	L	M
AI, 3-D, robotics	*	M	H	L	M
Dimension	Units/time	$/aggregate unit	$/system	$/product type	$/unit of product

4.4 Omni channel supply chain

Since the 2000s, e-commerce has grown at very rapid speed in the business-to-consumer (B2C) supply chain. The suppliers position themselves among pure e-commerce, traditional stores, or some hybrid of the two. In the United States, most suppliers have both an online and brick-and-mortar presence. Apple lets customers experience their products in Apple stores. Home Depot, Walmart, Macy's, and Barnes & Noble have both e-commerce and stores. Even Amazon has extended its boots on the ground to the end-consumers through its own delivery and through outlets such as Whole Foods stores.

The traditional supply chain with brick-and-mortar stores provides the best user experience for physical goods, even in the era of virtual reality. The stores must be located in a high-cost real estate area where people want to go. It must have stock up of physical goods, and provide appealing displays with a lot of room for customers to shop comfortably and personnel to help. These are very expensive. The worst part of this is that the great customer experience provided at high cost can be shared with others, including online retailers. You can consider this when you purchase your textbook.

The online services enjoy the huge benefits of economies of scale due to very low marginal costs. The services are information rich. They allow the user to search, to compare, to share experiences, to link to vast amount of physical inventories stored at low cost, to track, and to complete the financial transactions. Due to these benefits, the online services attract many individual consumers to shop for their own needs. The drawback of online services for physical goods is that the customer cannot experience the physical goods before they reach the customer's hands. Imagine a fashion product, where the fit, feel, and color are so important. Most online purchases provide good return policies. Some even send multiple products to let the consumer return the ones not selected.

The challenges in the omni channel supply chain are to share the infrastructure and operations to enjoy economies of scope between the online fulfillment and store deliveries. The supply chain towards the last part of the chain through the store is shown in the top part of Figure 8.5, and through the online purchasing in the bottom part of the figure. The store orders normally include

many line items and with many items in each line. The delivery is often once a day or once in a few days. The loads are normally in pallets, mixed pallets, cases, and some mixed cases. In contrast, the average online consumer order consists of fewer than two lines, and fewer than two items per line. The consumer-accepted lead time is very short. In the United States, from the time when an order is completed to the time when the order is ready to ship, is normally less than one day and often just a few hours. These fundamental differences in the orders and destinations make it difficult to scope up.

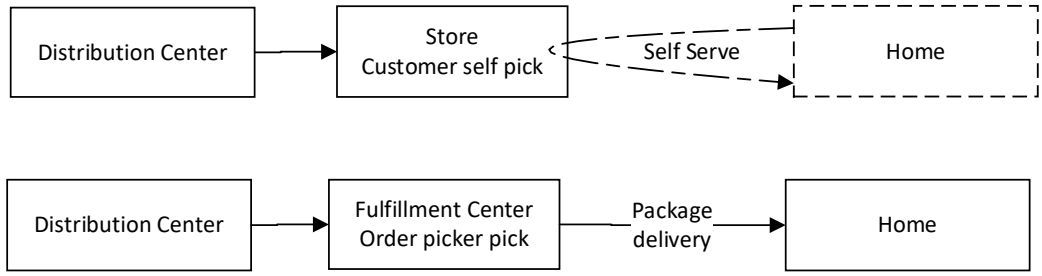


Figure 8.5 The last steps in the supply chain through stores and online delivery.

Amazon mainly has only one type of chain. Walmart has both types of chains. Walmart initially ran two separate chains, one for the stores and the other for online order fulfillment. However, Walmart found many ways to enjoy more and more scope effect between two chains. Although Amazon enjoys many e-commerce benefits, including using Walmart or other stores as Amazon's experience centers, many brick-and-mortar stores found way to compete. The stores also have some advantages of their own. Big chains already have many buildings close to customers while Amazon invests to build more and more buildings. As shown in Figure 8.5, online retailers such as Amazon will incur order picking, and most likely part of the shipping and handling costs that traditional stores do not incur.

5. ECONOMIES OF SCOPE IN PRODUCTS

5.1 Mass customization

The incremental cost of product scope can be very low, especially when the differences can be achieved digitally. Today, many products or services come with a quick-start manual, and with links to websites for further details. The websites can provide much richer information with illustrations or video instructions, much more effective than thick pamphlets in multiple languages. This is especially beneficial in Europe, where many languages are used in different countries.

HP coined the term "mass customization" in 1990s for its supply of printers in Europe. The combination of online service and computer interface in the product allows a single product to serve many markets, or "mass customization." Mass customization can simplify the management

of the supply chain, reduce inventory, and increase the ability for hedge risk. More later on this part.

Example

Different countries in Europe use different languages and different power outlet sockets. Consider an appliance marketed in Europe that covers eight languages and five types of three-prong power outlet sockets. Assume there are 25 different combinations. At one extreme, the supplier can supply 25 different product types, one for each country. At the other extreme, the supplier can supply a single product that can be used in all countries. Please

1. Develop two strategies to achieve the second strategy and discuss pros and cons of your strategy.
2. Compare the average inventory level without safety stock between two extreme solutions.

Solutions

1. The firm can

a. Adopt postponement of differentiation. The supplier can postpone the installation of language labels onto the appliance and power cables until a customer order is received. Before that point, the appliance is generic. The installation process of low-cost labels and cables can be done at the distribution center. In this strategy, the customer will get exactly what he or she ordered. However, the store can only serve as an experience center. The customer cannot purchase the product and bring it home because the store does not carry inventories. Light assembly at the DC can also cause delay.

b. Mass customization. Design the products with only digital displays and include a universal power adapter in the package, or all 5 types of low-cost power adaptors. The digital displays can support each of the eight languages. In this design, all appliances are the same but can be setup to satisfy each customer's demand. The customer can pick up the generic product at the store without waiting.

2. The inventory comparison is similar to the SKU reduction example in Aldi versus Walmart. Assume the demand for each type is D, identical among 25 types.

a. The total average inventory required for the expected demand using dedicated products is

$$\overline{I} \approx 25\frac{EOQ}{2} = \frac{25}{2}\sqrt{\frac{2KD}{h}} = 25\sqrt{\frac{KD}{2h}}$$

b. The total average inventory required for the appliance itself in the postponement of differentiation strategy and in complete package in the mass customization strategy is

$$\overline{I} \approx \frac{EOQ}{2} = \frac{1}{2}\sqrt{\frac{2K25D}{h}} = 5\sqrt{\frac{KD}{2h}}$$

In the postponement of differentiation, the DC must also carry inventory on low-cost labels and different power adapters with averages the same as in the dedicated products. In the mass customization, the universal power adapter or multiple adapters cost more.

This is only 20% of the other strategy! In reality, there will also be reduction in safety stock due to pooling effect, to be covered in the chapter on uncertainty.

5.2 Personalization

The computers, sensors, smart devices, and internet of things has already allowed firms to provide personalized product and services. The personalization can be based on our preferences explicitly expressed, such as putting initials on bags or cell phone cases, your photo on a mug or T-shirt, and so on. These systems can also detect our preferences from what we see, how we click, where we surf, how we talk, what we post, and provided services or relevant information. Research is being conducted to embed intelligent devices into our body to learn our thoughts, and collect biological information in our body. These can provide not just commercial products and services but personalized medicine.

6. SHARING ECONOMY

"Sharing economy" is a new term that came around during the great recession in 2008. It is often associated with Uber, Lyft, or Airbnb. As we have discussed throughout this chapter, there are many advantages to sharing things. "Sharing economy" refers to peer-to-peer (P2P) platforms that allow the clients and suppliers to match the supply and demand for short-term transactions, such as cars, bikes, rooms, tools, or freelance services. Some domains have faced more challenges than others.

The sharing economy allows better use of the idle capacities in terms of rooms, cars plus drivers, money, tools, and professional services. It also allows people to make income when there are needs. It has the potential to reduce the cost for the user and provide better utilization of goods and human resources.

The "sharing" also refers to sharing in time. Time-sharing allows the resources to be shared by different users to improve the utilization. If you remove the time-sharing, eBay, started in 1995, also provides a sharing economy. If resale is involved after use, it becomes a true sharing economy. Amazon started to allow third-party vendors to see through its platforms in 1999, called Amazon Marketplace. The vendors and buyers share the amazon P2P platform to complete the buying and selling.

7. ECONOMIES OF SHARING IN BARILLA CASE

Barilla's 200 varieties of dry pastas share a few types of rolled sheets on limited numbers of production lines built in different times in their four plants. The 470 pasta packages share these 200 pastas at a given time. Barilla's merchandizing analyzes the market to limit the package types to 800 total, such as for promotions or different motifs between the north and south. The reduction in the assortment with somewhat different requirement has to share limited product packages. However, the reduction can greatly reduce the supply chain complexity and cost. Barilla maximizes the sharing of raw materials, facilities, production lines, and personnel. In the Barilla-paid truckload shipment, Barilla allows the customer to share different pastas in the same load,

although this adds complexity in order picking. By allowing 200 products completed in a much lower number of production lines, Barilla also enjoys the economies of scope. The sharing along all these dimensions allows Barilla to stay competitive with low-cost and high-quality products.

There is plenty of room for improvement. Is the product assortment optimal? Can the setup cost and time be reduced to improve the economies of scope, to reduce inventory, and improve response time? Are there opportunities to postpone product differentiation such as packaging?

8. ECONOMIES OF SHARING FOR HUMAN NEEDS

Sharing among a limited product or service assortment can reduce cost or increase resource utilization. Please count how many products and services are now available on our smart phones. It can also reduce the level of service to human needs, such as a limited variety of offerings; or waste, such as being too big or having too much functionality. The reduction of assortment can also reduce life satisfaction, the ultimate goal for human beings. The sharing of resources can reduce the cost or increase the utilization to supply for human needs. The sharing of resources also has potential to reduce the environmental footprint to supply a variety of products and services. The computer, information technology, and artificial intelligence have reduced the cost of flexible systems with greater and greater scope, reducing the need for the variety of resources. It also increases the needs for the educated labor force but reduces the need for the unskilled labor force. Together, it has the potential to increase inequality.

9. UNCERTAINTY, ITS COST AND REDUCTION

We love certainty. However, those who embrace and plan for the inevitable uncertainty thrive. The demand, market, human performance, weather, pandemic, natural disasters, and potential interruptions are usually uncertain. Uncertainty increases risks and cost. A supply-chain professional must make decisions against these uncertainties. However, one thing is certain: if your decision is not spot on, which is almost certain, then there will be penalty due to either too much or too little.

To be competitive, supply-chain professionals should understand uncertainty, model uncertainty, model the costs associated with too much or too little, and design strategies and solutions to minimize the negative impact of uncertainty.

One approach is to capture uncertainty with probability distributions, either continuous or discrete, and model the expected cost of being too much or too little, and find the optimum solution.

1. MEASURES OF UNCERTAINTY

There are many ways to measure uncertainty. Most supply-chain professionals know the concepts of range, spread, standard deviation, and margin of error. Some may also know variance, coefficient of variation, *p*-value, and confidence interval. Some of these measures have a single parameter while some have two parameters. Some single-parameter measures are absolute, such as standard deviation or variance; while others are relative such as coefficient of variation or *p*-value. The two-parameter measures normally include both absolute and relative measures such as range or confidence interval. Some are linear, such as standard deviation, while others are not, such as variance. Some require a significant background in probability and statistics to understand, such as *p*-value and confidence interval, while others require very little background, such as range.

If you are allowed to use a single, linear, and relatively simple parameter to provide a rough idea about the level of uncertainty of a quantity—such as the number of people who show up at a meeting, the weekly demand, the probability a container will arrive in two weeks, or the probability that the surgery will finish in one hour—would you be interested in a relative or absolute measure? Most people would want a relative measure, such as percentage of the people informed to show up in the meeting, the relative range of the deviation of the demand, the percentage deviation from the two-week nominal shipping time, and the percentage deviation from the one-hour nominal surgery time.

1.1 Coefficient of variation

Due to its simplicity, linearity, and relative nature, the coefficient of variation has been used more and more widely in business. Please compare pairs of uncertainty of number of bottles of one 16-ounce bottled drinking water sold in different sites for different time durations shown in the table below.

Site \ Time	Day			Month		
Measures	μ	σ	CV	μ	σ	CV
Vending	5	5	1	150	30	0.2
Plant	100,000	10,000	0.1	3,000,000	51,000	0.017

For each site/time combination, the average sales, the standard deviation of sales, and the coefficient of variations are shown. All three independent parameters are important in the design and operation of a supply chain. The average is useful to get the magnitude of information. Standard deviation is useful to get the absolute spread in the magnitude. The daily data is useful for restocking and production planning, while the monthly data provides data for financial analysis.

The standard deviation on one vending machine in a day will be the lowest, while the standard deviation in a bottling plant in a month will be the highest due to the huge difference in the averages. However, the average is not related to uncertainty. Parameter pairs such as range or confidence interval would provide measure of uncertainty. However, both require two parameters, and the confidence interval requires special training.

Here, the coefficient of variation (CV) comes in. CV is a single, linear, and relative measure of uncertainty and is easy to explain, and widely adopted in many domains. It is more and more familiar to many working in the supply chain. The coefficient of variation is defined as the ratio of standard deviation over the mean, or

$$CV = \frac{\sigma}{\mu}$$

Or sometimes, the squared coefficient of variation is used, which is not linear.

$$SCV = \frac{VAR}{\mu^2}$$

Example

Using this definition, we can find the relative variability for the number of cokes at the beginning of this section.

$$CV_{machine-day} = \frac{5}{5} = 1$$

$$CV_{plant-day} = \frac{10{,}000}{100{,}000} = 0.1$$

$$CV_{machine-month} = \frac{30}{150} = 0.2$$

Intuitively, the daily sales from a vending machine can vary a lot due to the activities in the area, the weather, type of people visited, and many other factors. Some days nothing, but occasionally over 10 because a large group played a game in the area on a hot day. Since the capacity in the machine in one slot is 12, it ran out of stock in the afternoon.

A plant supplies many stores, restaurants, vending machines, and other places. A person who buys the water at a vending machine may not buy from the store. The uncertainties compensate each other, and therefore reduce the relative uncertainty. This is pooling effect among channels.

The same machine can have high sales and low sales on different days in a month. The high and lows can reduce the relative uncertainty when considered together. This is pooling effect over time. More on the strategy on pooling in uncertainty reduction later.

1.2 Coefficient of variation of some common random variables

Let us consider the range of some commonly used random variables. Typical quantities in a supply chain, such as demand, arrivals, processing times, and quantity supplied, do not take negative values. For these quantities, the coefficients of variations are only valid within some range, shown in CV_{max} column in Table 9.1. Some quantities can have negative values, such as back-ordered stocks. Back-ordered stock refers to the situation when there is stockout and the demand will be satisfied by future stocks, as opposed to lost sales. Even in that situation, the back-ordered stocks are different from positive stock levels and must be treated differently and involve more complex modeling.

Table 9.1 CV for some commonly used random variables in supply chain.

	Pdf	$\boldsymbol{\mu}$	$\boldsymbol{\sigma}$	***CV***	$\boldsymbol{CV_{practical}}$	**Use**
Exponential Continuous $Exp(\lambda)$	$\lambda e^{-\lambda x}$	λ^{-1}	λ^{-1}	1	1	Time between arrivals, time between interruptions, some service times
Poisson discrete $Pois(\lambda t)$	$\frac{(\lambda t)^k e^{-\lambda t}}{k!}$	λt	$\sqrt{\lambda t}$	$\frac{1}{\sqrt{\lambda t}}$	$(0, \infty)$	Number of arrivals or interruptions during time *t*
Normal continuous $N(\mu, \sigma^2)$	$\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ	σ	$\frac{\sigma}{\mu}$	< 1*	Sum of many random variables

	Pdf	$\boldsymbol{\mu}$	$\boldsymbol{\sigma}$	***CV***	$\boldsymbol{CV_{practical}}$	**Use**
Uniform continuous $U(a,b)$	$\begin{cases}\frac{1}{b-a}, x \in [a,b] \\ 0, \quad otherwise\end{cases}$	$\frac{a+b}{2}$	$\frac{1}{\sqrt{12}}(b-a)$	$\frac{1}{\sqrt{3}}\frac{b-a}{b+a}$	0.577	Travel distances in warehouse, etc.

* If normal distribution is used to approximate some positive quantity, a $CV \gg 0.333$ implies a large negative tail ignored. If $CV = 0.333$, the $\Phi(-3)$ is negligible. If $CV = 0.5$, $\Phi(-2) = 0.023$, which is significant but may be acceptable. At $CV = 1$, $\Phi(-1) = 0.16$, probably not acceptable. Therefore, if you use normal distribution, CV cannot be too high.

For exponential distribution, $CV_{Exponential} \equiv 1$. It often provides good approximation in uncontrolled interarrival times. It sometimes provides good approximation for certain service times when most service times are very short, some near zero, while some can be very long.

The Poisson distribution refers to the counting of number of occurrences with time between occurrences exponentially distributed. It provides good approximation for the number of uncontrolled arrivals, such as the demand rate for a given time interval, number of interruptions in a time interval, or number of emails received in an hour or in a day. The Poisson distribution is the only commonly used random variable listed in the table that can have a high CV. Note that the $CV_{Poisson} \propto \frac{1}{\sqrt{\lambda t}}$, or it increases with the reduction in its constant parameter λ and the time interval you count the arrivals, such as per hour, day, or month. It is possible $\lambda t \ll 1$ when either λ or t is very low. For the same average inter-arrival time, the longer the time interval to count, the lower the relative uncertainty; this is pooling over time!

1.3 Ranges of CVs

Certain activities tend to have high levels of uncertainty while others have less. For example, when an automated injection-molding machine is producing the same parts in an assembly line, the variability in processing time is very low. This is true for many types of mechanized production machines. On the other hand, the service time need for a serviceperson to fix the injection-molding machine that stopped working can be highly variable. Hopps and Spearman (2011) recommended three levels of variabilities, $0 < CV_L < 0.75$, $0.75 < CV_M < 1.33$, and $1.33 < CV_H$, shown in Figure 9.1.

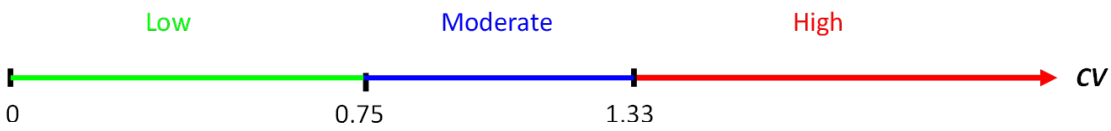


Figure 9.1. The levels of uncertainty or variability.

Each range is quite wide. The processing time of identical tasks at a workstation tends to be low. However, the variation in manual processing is higher than that of machines. When the task varies, the variation will be even be higher, such as a worker installing different options in an automobile assembly line, or an order-picker selecting an item in a zone. Note in such operations, uniform distribution is already highly variable. Even with uniform distribution, without negative values, the maximum CV is 0.577!

The time it takes to repair something in the supply chain tends to have higher relative variability. You may have this experience with your car. In a supply chain, getting the injection-molding machine to work again can be a simple fix, such as an adjustment or fixing the raw material supply. This can be done in seconds or minutes. But the problem could also be a damaged heater, and service in that case could take hours. The same thing applies to delivery. A delivery truck is delivering a truckload of goods to a client but interrupted along the way. It can be a simple fix by a mechanic. It can also be a major problem that cannot be fixed. The firm can find a substitute truck to accomplish the job. But it has to wait to find the truck, get the truck to the location, transfer the load to the replacement truck, and so on, which will take a very long time. For this type of service time with very high variabilities, exponential distribution may be useful in modeling. Its coefficient of variation is 1, around the middle in moderate variability.

What are the processes that lead to high variabilities? The coefficient of variation increases with standard deviation and inverse of the mean. To achieve high CV, you would need a low mean with high absolute variability.

Example

The weekly sales for a gardening book in a bookstore for the last 10 weeks are 0, 0, 0, 1, 0, 0, 0, 0, 0, 0. The annual sales in the last five years are 2, 5, 3, 1, 4 books each year.

1. What is the coefficient of variation for weekly demand?
2. What is the coefficient of variation for one year?
3. What can you observe from this simple example?

Solutions

1. We can find the mean and standard deviations using the general equations. You can also verify with other tools.

$$\mu_{week} = \frac{1}{10} = 0.1$$

$$\sigma_{week} = \sqrt{E[X^2] - (E[X])^2} = \sqrt{\frac{1}{10} - \left(\frac{1}{10}\right)^2} = \sqrt{0.09} = 0.3$$

$$CV = 3$$

This is very high. It is very difficult to manage. Only the Poisson process can accommodate.

2. Similarly,

$\mu_{year} = \frac{15}{5} = 3$

$\sigma_{year} = \sqrt{E[X^2] - (E[X])^2} = \sqrt{\frac{4+25+9+1+16}{5} - 3^2} = \sqrt{2} = 1.414$

$CV_{year} = \frac{1.414}{3} = 0.47$

This is in the middle of low range. Normal distribution is possible.

3. The magnitude increases with the difference between the mean and individual sales and decreases with the averages.
    a. When most of the values are low, with zero being the lowest if not negative, the mean is low. If you couple that with large differences between the mean and one or few large values, the standard deviation is high. Together, the *CV* becomes high.
    b. If the single weekly sale is five by someone who wants to give the book as gifts, what is *CV*? In what situation *CV* can be higher than three?
    c. The *CV* reduces as you increase the length of window accounting. You can see it in the Poisson distribution.

The infrequent number of occurrences during the observation period, or exceptional events such as major interruptions, can be highly variable. Consider how many days before a car breaks down? This is true in manufacturing, delivery, order-picking, and other parts of the supply chain, too. Such types of interruptions can be most damaging. Imagine you drive to a final exam, to an interview, or to a client site to solve an emergency. There are ways to reduce the probability of such happenings: preventative maintenance.

2. THE COST OF UNCERTAINTY

In the supply chain, over-supply requires space for storage, tied up capital, subjected to damage, price drop, obsolescence, or other penalties. The undersupply at best would lead to back order and unhappy customers if the customer can tolerate the delay. In worse situations, it can mean lost sales. This is why forecasting is so important. However, if inherent uncertainty exists, forecast results will always be uncertain. A better forecast only means the removal of the systemic variations such as trend of seasonality. The potential cost associated with over-supply or under-supply, in schedule or in capacity, and associated best strategies depend on many situations. There are many tools and models such as the newsvendor model, Markov decision process, and stochastic optimization that can be used to find the best strategies.

The simplest model, and probably one of the most useful models, is the newsvendor model. We will use this simple model to illustrate the cost of uncertainty, and the strategies.

The newsvendor model applies more widely than minimizing the cost of selling perishable items. For example, in capacity planning, you can either plan too much or too little. Too much capacity would mean idle capacity in facilities and personnel. Too little capacity would mean unsatisfied

demand, overtime operations, and so forth. The capacity can be in space, number of personnel, number of pieces of equipment, size of a fleet, and rent or build.

In a supply chain, you may need to decide based on an arrival of a shipment, say in two weeks. However, two weeks is only a nominal time. In practice, it can be earlier or later. When it is too early, there is a cost associated with it, just as when it is too late.

Operating rooms are expensive resources. They require highly skilled personnel to operate. However, the surgery times are highly uncertain, depending on many factors plus potential complications. If the scheduler schedules the next patient while the current patient is not finished, the next patient, and all the associated personnel and resources, have to wait. They will occupy the expensive preparation space and make the personnel unproductive. If the wait time is too long, certain preparation work may have to be done again. If the current operation is finished before the next patient is scheduled for operation, the expensive operating room must stay idle.

2.1 Newsvendor model

The newsvendor model can be used to make a single decision to trade between too much and too little against a single uncertain quantity, given that there are known probability functions and known costs proportional to the amount of too much or too little. The name refers to the decision a newspaper vendor has to make each morning: how many papers to acquire in the morning against an uncertain number of newspapers to be sold. If the vendor acquires too many copies, there will be worthless newspapers left at night. On the other hand, if the vendor acquired too few copies, he or she may run out of the newspapers too early and lose potential profits.

2.1.1 Parameters and assumptions

The parameters and assumptions of the newsvendor model are given below.

1. $f(x), F(x), p(x), P(x)$. The random quantity can be modelled with the probability density and cumulative functions for continuous variables or their discrete counterparts. In the newsvendor model, this will be the demand for newspapers. In the operating room, this will be the operating room ready time. In the capacity planning problem, this will be capacity to be needed.
2. Q: The decision variable. In the newsvendor problem, this is the number copies of newspapers to be acquired in the morning. In the operating room schedule, it will be the next patient scheduled start time. In the capacity planning problem, this will be the planned capacity.
3. c_O: The constant overage cost per unit per decision. In the newsvendor model, this will be the cost of the loss from the extra copies of papers not sold at the end of the day. For the operating room, this will be the cost of due to patient wait.
4. c_u: The constant underage cost per unit per decision cycle. For the newsvendor problem, this will be the loss of profit, or the selling price – acquisition cost.
5. $C(Q^*)$: The minimized expected total cost incurred by optimum decision Q^*.

Compared to other models in supply chains and economics, the assumptions in the newsvendor model are more practical and generic. Therefore, it is quite powerful. At least it helps people to think qualitatively if the distribution and parameters are not available.

2.1.2 Newsvendor model formulation

For a continuous-demand probability density distribution $f(x)$, shown in Figure 9.2, and the order quantity of Q, we can find the expected cost for the decision in verbal form as

$$C(Q) = c_o E[\text{number of units over}] + c_u E[\text{number of units under}]$$

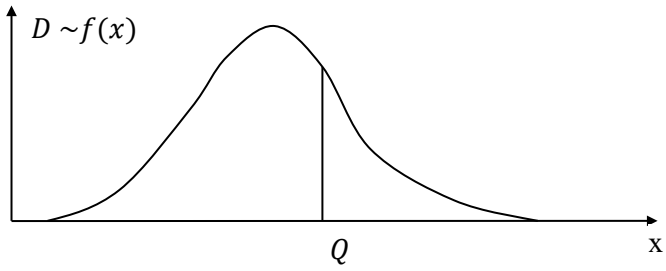


Figure 9.2. The newsvendor problem graphical illustration.

The number of units over is when the order quantity is more than the demand, or the total area from x to Q under the curve, shown in the first term below. The number of units under is when the demand is higher than the order quantity, or from x to ∞, shown in the second term below.

$$C(Q) = c_o \int_{-\infty}^{Q} (Q - x) f(x) dx + c_u \int_{Q}^{\infty} (x - Q) f(x) dx$$

For discrete distribution, we have the following. Note that for any given $Q - Q_i$, $Q - Q_i = 0$. We do not have to include it in the range of summations.

$$C(Q) = c_o \sum_{Q > Q_i} (Q - Q_i) p(Q_i) + c_u \sum_{Q < Q_i} (Q_i - Q) p(Q_i)$$

For the continuous case, we can apply first-order condition to find the optimum solution Q^*. Since Q is also in the limits, you need to apply the Leibniz rule.

$$\frac{d}{dQ} \int_{a(Q)}^{b(Q)} f(Q, x) dx = \int_{a(Q)}^{b(Q)} \frac{\partial}{\partial Q} f(Q, x) dx + f(b(Q), Q) \frac{\partial b(Q)}{\partial Q} - f(a(Q), Q) \frac{\partial a(Q)}{\partial Q}$$

$$\frac{dC(Q)}{dQ} = c_o \int_{-\infty}^{Q} \frac{\partial(Q-x)f(x)}{\partial Q} dx + (Q-Q)f(Q)\frac{dQ}{dQ} - (-\infty - Q)f(-\infty)\frac{d(-\infty)}{dQ} + c_u \int_{Q}^{\infty} \frac{\partial(x-Q)f(x)}{\partial(Q)} dx + 0 + 0$$

$$\frac{dC(Q)}{dQ} = c_o \int_{-\infty}^{Q} 1 * f(x)dx + c_u \int_{Q}^{\infty} (-1)f(x)dx = c_o F(Q) - c_u[1 - F(Q)] = 0$$

We can find the condition for the optimum Q^*:

$$F(Q^*) = \frac{c_u}{c_u + c_o}$$

If the probability function is the discrete case, it can be proved that the solution is the quantity that yields the cumulative mass function greater than or equal to the critical ratio, or

$$P(Q^*) \geq \frac{c_u}{c_u + c_o}$$

In a supply chain, most quantities are in integers. A continuous distribution is an approximation for the uncertainty. If the result can only take integer values—such as number of items to order, scheduled arrival time on the hour, number of rooms to reserve, number of cash registers to order—you should round up the result.

2.1.3 The solution for the decision variable in newsvendor problems

This is an inexplicit solution in the form of $F(Q^*) = \cdots$. The left-hand side represents the cumulative probability function at the optimum decision quantity. The right-hand side is a ratio of the underage cost over the sum of underage and overage costs, or a constant called critical ratio, and depends only on the cost ratio with values between (0, 1). We do not have to worry with the case of 0 or 1, because the decision is obvious when there is no cost for underage or overage.

$$0 < \frac{c_u}{c_u + c_o}(Critical\ Ratio) < 1$$

If the underage cost is very low, the critical ratio approaches zero. If the overage cost is very low, the critical ratio approaches 1. As the cumulative probability function, $F(Q)$ is continuous, non-decreasing function going 0 to 1. At some point when $x = Q^*$, the total cost is minimized. The larger the critical ratio, the larger the Q^*. This makes sense because the higher the underage cost,

the higher you should set the quantity; and the higher the overage cost, the lower you should set the quantity. When $c_u = c_o$, the costs of too much or too little are the same. The critical ratio is equal to 0.5. That means that the minimum cost occurs at the median!

2.1.4 Minimum cost in the newsvendor model

Separate the terms from the model, and applying the properties of probability distribution functions, $\int_{-\infty}^{Q} f(x)dx = F(Q)$ and $\int_{Q}^{\infty} f(x)dx = [1 - F(Q)]$, we have

$$C(Q) = c_o QF(Q) - c_o \int_{-\infty}^{Q} xf(x)dx + c_u \int_{Q}^{\infty} xf(x)dx - c_u Q(1 - F(Q))$$

If we can solve $\int_{-\infty}^{Q} xf(x)dx$, we can find the cost in closed form. If demand is normal, we can express it in the standard normal form:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} = \frac{1}{\sigma}\varphi\left(\frac{x-\mu}{\sigma}\right)$$

Let $x = \mu + \sigma z, dx = \sigma dz$, when $x = -\infty, z = -\infty$, when $x = Q, z = \frac{Q-\mu}{\sigma}$. Then,

$$\int_{-\infty}^{Q} xf(x)dx = \int_{-\infty}^{\frac{x-\mu}{\sigma}} (\mu - \sigma z)\frac{1}{\sigma}\phi(z)\sigma dz = \mu\Phi\left(\frac{x-\mu}{\sigma}\right) + \sigma\int_{-\infty}^{\frac{x-\mu}{\sigma}} z\phi(z)dz$$

Then,

$$\int_{-\infty}^{Q} xf(x)dx = \mu F(Q) - \sigma\phi\left(\frac{x-\mu}{\sigma}\right) = \mu F(Q) - \sigma\phi(z)$$

Similarly,

$$\int_{Q}^{\infty} xf(x)dx = \int_{-\infty}^{\infty} xf(x)dx - \int_{-\infty}^{Q} xf(x)dx = \mu - (\mu F(Q) - \sigma\varphi(z))$$

Putting all back to the original model, we have

$$C(Q) = c_o QF(Q) - c_0\mu F(Q) + c_0\sigma\varphi(z) + c_u\mu F(Q) + c_u\sigma\varphi(z) - c_0 Q(1 - F(Q)$$

Factoring out the underage and overage cost, we have

$$C(Q) = c_o[(Q-\mu)F(Q) + \sigma\varphi(z)] + c_u[(\mu - Q)(1-F(Q)) + \sigma\varphi(z)]$$

At $Q = Q^*$, we find the minimum cost, which is the cost of uncertainty! If there is no uncertainty, this cost is always zero. Of course, if you select $Q \neq Q^*$, the resultant cost includes both the cost of uncertainty and error in the quantity selection.

2.2 Applications of the newsvendor model in the supply chain

In a supply chain, the newsvendor is useful in many practical situations.

1. Inventory problems
    a. One-time decisions as with newspapers, other perishable goods, seasonal goods, or high fashion.
    b. Repeated decisions in which the overage can be linked to the next decision such as periodic deliveries, daily, weekly, monthly, and so on. The overage goods can be used and modeled in the next decision cycle.
2. The scheduling problem in services, such as in the operating room. This can occur in other too-early or too-late types of problems.
3. Capacity design for the size of a facility, number of personnel, number of equipment such as cash registers, number of trucks, number of rooms, size of the storage area.
4. Determine the resources in contact tracing in pandemics. When the scope is too large, there is a cost associated with the tracing effort. When the scope is too small, there is the cost of potential spread.
5. Other problems when the problem fit the general assumptions in 2.1.1.

2.3 Some notation in simple inventory models

Inventory is an important issue in supply chains. Let us first consider some simple inventory-related problems. We may encounter some inventory terms in a supply-chain context. You may find the same term called differently by different groups or texts. The following defines the intended meanings in this class.

2.3.1 Some definitions of inventory problems

The definitions are given below. Each definition can be given or a decision. Each definition can be deterministic or probabilistic.

- Inventory review time: The time interval inventory level is reviewed to determine if it is time to order, and if yes, how much to order. Inventory control policies are often divided into continuous review and periodic review policies. In modern systems with warehousing management systems (WMSs), the inventory level is updated whenever a bar code (or RFID) is scanned. That is why when you order from Home Depot or Amazon, they show how many items are in stock. This makes it almost real time. In traditional systems, however, inventory positions are updated only at certain points when the inventory is tallied, rather infrequently, say, every week.
- Inventory reorder time: The time an order is placed.

- Replenishment: The activity to restock the inventory.
- Order lead time (LT): The time from when the order is placed to when the goods are replenished.
- Inventory order cycle time (CT): The average time between replenishments.
- Production or order quantity (Q): This can be a fixed amount such as a unit load: a truckload, a container, a pallet. This can also be a variable depending on the current inventory level and the target level to stock up to.
- Reorder point R: An inventory level below which an order is placed.
- Window of uncertainty (W): The duration in which the demand realized is uncertain. In our simple models, $W = LT + CT$.
- Order-up-to level: The expected level of inventory if the order arrives at the time the order is placed.
- Safety stock (SS): The expected inventory level right before the replenishment.
- Cycle stock (CS): The average inventory sold in one ordering cycle.
- Stockout: When the inventory reaches zero. During stockout, there may or may not be orders. If there is an order, and the customer will wait until the next order arrives, it is considered back-ordered. If the customer orders from other suppliers, the stockout leads to lost sales.
- Back order: During stockout, if the customer will wait for the new stock to arrive. The penalty can be the cost of exception handling and the cost of loss of good will.
- Lost sales: During stockout, if the customer does not want the product—or worse, takes the business to others, such as a competitor.

2.3.2 One-time decision newsvendor application

This is what the original model intended, similar to newsvendor problems. Let us look at an example.

Example

A dealer orders and sells windows. The retail price is $300 per window. The wholesale price the dealer pays the manufacturer is $200 per window. The unsold units by the late fall will be sold to the secondary market at $60 per window. The dealer estimates the demand in the year can be approximated by a normal distribution $N(300, 100^2)$ per year. The production cost per window is $120. The setup cost for each order is $1,000.

1. How much should the dealer order from the manufacturer?
2. What are the expected numbers over and under?
3. What is the expected cost per order?
4. What does this cost mean?

Solutions

Assume that the demand forecast is accurate, the costs estimates are good, and the buyer in the secondary market is reliable with the price given.

1. How much should the dealer order from the manufacturer?

The overage cost is the difference between what the dealer paid to the manufacturer and recovered from the secondary market.

$$c_o = 200 - 60 = 140$$

The underage cost is the loss of potential profit.

$$c_u = 300 = 200 = 100$$

The critical ratio

$$F(Q^*) = \frac{100}{100 + 140} = 0.4167 < 0.5$$

This value is not in the table on the formula sheet. Some transformation is needed

$$z = \Phi^{-1}(0.417) = -\Phi^{-1}(1 - 0.417) = -0.21$$

$$\phi(-0.21) = \phi(0.21) = 0.39$$

If not during a test, you can use other ways to find these values. In Excel, NORMSINV(.) will return −0.21034. You can round it off to −0.21, You can also use NORMDIST(., 0, 1, False) to get 0.390214, which you can off to 0.390.
The order quantity is

$$Q^* = 300 + 100 * (-0.21) = 279$$

This is 21 units less than the expected demand. It can be considered negative safety stock.

2. The expected number of windows to be sold to the secondary market is

$$E(unsold) = (279 - 300) * 0.417 + 100 * 0.39 = 30.24$$

The expected number of windows short is

$$E(number\ short) = (300 - 279)(1 - 0.417) + 100 * 0.39 = 51.24$$

3. The cost associated with the order size of 279 is

$$C(279) = 140 * 30.24 + 100 * 51.24 = 4{,}237 + 5{,}124 = \$9{,}361$$

With the expected overage cost of $4,327 and underage cost of $5,124. The two values are similar because the costs of 140 and 100 are not too different. If the underage cost is much higher than the overage cost, such as in the fashion industry, the critical ratio will be very high. On the other hand, Costco assigns very low underage cost to seasonal goods, and the critical ratio is very low, which will lead to relatively higher expected shortage cost compared to overage cost but overall low magnitude due to the low unit cost assigned.

4. What does this cost mean?

To supply windows, there are many other costs involved. This model only captures one aspect of these costs: uncertainty in demand. If the dealer can somehow change the model, such as to order with much less uncertainty in demand, this cost can be reduced. More about this in later sections and chapters on the reduction in uncertainty.

2.3.3 The fixed replenishment cycle model

In a supply chain, many firms replenish inventories between fixed cycle times, such as every week. Graphically, the actual inventory level over time may look like that in Figure 9.3.

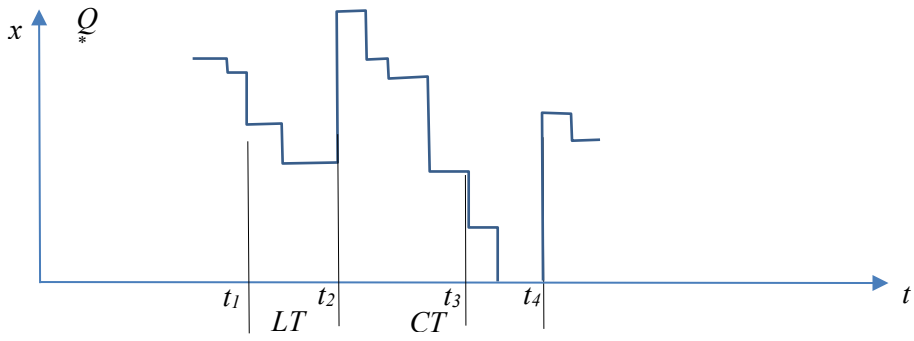


Figure 9.3 Illustrations of inventory levels over time for an SKU.

The horizontal axis is time. The vertical level is inventory levels. The stepped line indicates the actual inventory levels over time. An order is placed at t_1. The order arrives and replenishes the inventory at time t_2. At t_3, the second order is placed. Sometimes after t_3, two large orders depleted the entire inventory and stockout occurred. At t_4, the second order is replenished. However, this order did not consider the large inventory drops before its arrival.

The expected inventory levels can be modeled using piece-wise linear lines shown in Figure 9.4. At t_1, it is the time to place an order, with a desire that the current inventory, plus the new order, will reach the order up to quantity Q'. Right before t_2, the inventory is expected to drop to the

expected safety stock level *SS* during *LT*. At this time, the new order is replenished and the stock level will reach the optimum level Q^*. Another order is placed at t_3, and replenished at t_4.

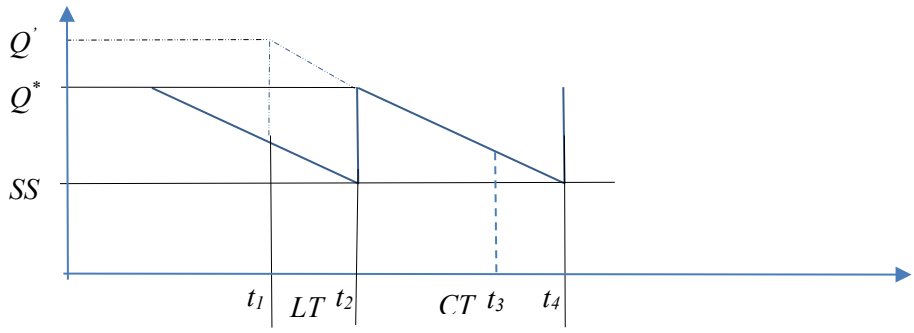


Figure 9.4. Model for the fixed ordering cycle problem.

Let us consider how the newsvendor model can be used to determine the amount to order. At time t_1, we need to find Q^* to trade off extra inventory or shortage at the time of t_4. The time window includes both LT and CT because the inventory position right before the second order arrives is determined by the order size placed at time t_1. Therefore, the window of uncertainty is $W = LT + CT$.

There are two parts in Q^*. It should include the expected inventory drop during the cycle time and the safety stock.

Example

The demand for a popular tool is $N(1000, 300^2)$ per week. The firm orders every three weeks. The order lead time is one week. The retail price is \$100, and the wholesale price is \$65. The stock out is lost sales. The holding cost is \$0.2 per tool per week. The loss of good is \$5 per unit per occurrence.

1. What should be the average order quantity?
2. What is the window of uncertainty, and demand during the window of uncertainty?
3. What are the overage and underage costs, and the critical ratio?
4. What are the order-up-to level and the safety stock?
5. The last order was replenished three weeks ago. The current inventory level is 1,100. How much should you order?
6. What if you have to order by pallet with 100 tools each?
7. What is the cost of uncertainty of this tool per year, assuming 52 weeks per year?
8. What is inventory-related cost for the cycle stock per year?
9. What can be done to reduce inventory-related cost?

Assumptions

1. The demand for the same tool does not change during the planning horizon. If the model changes, the demand can change, especially during the transition.

2. The demand in each week is iid. In practice, the demand between weeks can be correlated.
3. The above assumption also applies to future demand, which is independent of prior demand.
4. The three-week order cycle is near optimum to balance setup cost and carrying cost.

Solution

1. The average order quantity should be equal to the average demand in three weeks, or

$$\bar{Q} = 3 * 1000 = 3000$$

2. The window of uncertainty is the sum of the lead time and order-cycle time, or

$$W = 1 + 3 = 4$$

The standard deviation, with iid assumption of weekly demand, is

$$\sigma_W = \sqrt{4 * 300^2} = 600$$

3. The overage cost is the holding cost during the entire window of uncertainty, because we make the decision every three weeks, with four weeks of uncertainty duration.

$c_o = h = 0.2 * 4 = 0.8$ per tool per cycle.

The cost of shortage or underage is the loss of profit opportunity plus the \$5 loss of good will, which account for the customer's unhappiness about the supplier.

$c_u = 100 - 65 + 5 = 40$ per tool shortage per cycle.

The critical ratio is

$$F(Q^*) = \frac{40}{40+0.8} = 0.980$$

Using the normal table, $z = 2.05$. $\phi(2.05) = 0.049$. The latter is useful to find the cost.

If not during a test, you can use other ways also. For example, in Excel NORMSINV(.) will return 2.053749. You can round it to 2.05 or 2.054. You can also use NORMDIST(., 0, 1, False) to get 0.049, or round to 0.0488.

4. What are the order-up-to level and safety stock?

The order-up-to level is the average order quantity, or the demand in 3-week interval, plus the safety stock, or

$$Q^* = \mu + \sigma z = 3000 + 600 * 2.05 = 4{,}230$$

$$SS = \sigma z = Q^* - \mu = 1{,}230$$

5. Using the order-up-to level or safety stock, we can find the order size when the inventory at the time of the review, or two weeks after last order arrival.

The order quantity should be adjusted based on the current inventory position. Since the current inventory is 1100, which is 130 less than the expected safety stock level, we will order 130 more than the expected demand, or 3000 + 130 = 3130.

6. If you have to order by multiples of 100.

You should round up due to the ≥ sign in the critical ratio, or order 3200.

7. The cost per order cycle of three weeks at the above order-up-to level is

$$C(4230) = 0.8[(4230 - 3000)0.98 + 600 * 0.049] + 40[(3000 - 4230)(1 - 0.98) + 600 * 0.049]$$
$$= 0.8[(1230)0.98 + 29.4] + 40[(-1230)(0.02) + 29.4]$$
$$= 987.8 + 192.0 = 1{,}179.8$$

The cost of uncertainty per order cycle is approximately $1180 or in every three weeks. About 84% is attributed to the holding of safety stock, and 16% attributed to shortage or underage. The total inventory cost due to uncertainty is 1,179.8 * 52/3 = $20,450.56.

Note that this is only considered the cost of uncertainty alone. If there is no uncertainty, this cost is zero.

8. The total inventory-related cost also includes the cost due to cycle cost and setup cost, such as shipping.

The cost due to cycle stock is the average cycle stock times cost of holding, or 1500 * 0.2 * 52 = $15,600. If the average order size is near EOQ, the total inventory cost, including setup, is twice this, or $31,200. Therefore, the total inventory-related cost will be $20,451.+ $31,200 = $51,651 per year. Of which, about 40% is due to uncertainty!

9. Several things can be done to reduce inventory-related cost.

If you work with your client to reduce demand uncertainty or order lead time, you can reduce the window of uncertainty. Consider Barilla's case.

Barilla can work with the distributor to smooth out the demand. The distributors order from Barilla weekly. The average order lead time is 10 days, or 1.5 weeks, which ranges from 8 days

to 15 days. If Barilla can reduce the average lead time or better yet, reduce the order lead time variability (we will see later), the distributor can reduce the safety stock greatly. If Barilla can work with the distributor to reduce the order setup cost, the distributor may order more frequently, which will reduce both the window of uncertainty and the cost due to cycle stock. Although more frequent deliveries will reduce the economies of scale.

2.3.4 The reorder point model with fixed order quantity

Another simple and practical inventory policy is reorder point with fixed order quantity Q model. The decision variable is the reorder point R, shown in Figure 9.5. The inventory level is continuously monitored. Once the inventory level drops to the level R, say at time t_1 in the figure, a fixed order quantity Q is placed. The order quantity can be near EOQ, at the minimum order quantity (MOQ) or a pallet or a container. This order will be replenished at t_2 after LT. After that time, the inventory will drop. At t_3, the inventory drops down to R again, and another order of Q is placed. The order placed at t_1 will be replenished at t_4.

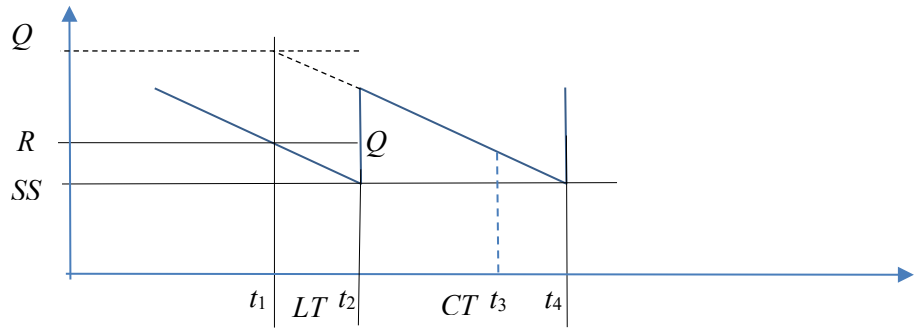


Figure 9.5. Model for the fixed ordering quantity reorder point problem.

Example

Let us consider the same example with a different policy. The demand ~ $N(1000, 300^2)$ per week. The firm orders one 20-foot container load each order. It contains 3000 tools. The order lead time is 1 week. The retail price is \$100 per tool, and the wholesale price is \$65. The stock out is lost sales. The holding cost is \$0.2 per tool per week. The loss of good will cost is \$5 per unit per occurrence.

1. Please find the reorder point.
2. The last order was replenished three weeks ago. The current inventory level is 1100. How much should you order?

Assumptions

1. Assumptions 1 – 3 as before.
2. The unit load of a container is near optimum quantity to balance the setup cost and carrying cost.

Solution

1. This is a reorder point, fixed-order quantity problem.

The quantity in a container is equal to the expected demand in three weeks. Therefore, on average, the order cycle is three weeks. Therefore, the expected costs and the critical ratio are the same as before. The difference is that the cycle time is no longer a constant.

$c_o = h = 0.2 * 4$ per tool for an average of 4-week during the expected window of uncertainty.

$c_u = 100 - 65 + 5 = 40$ per tool short per ordering cycle.

$F(Q^*) = \frac{40}{40+0.8} = 0.98$

The approximated results are the same as in the fixed replenishment cycle model. The problem is to find reorder point R which is the order up to level Q^*, show on Figure 9.5, minus the average demand during the 4-week window of uncertainty, or

$Q^* = \mu + \sigma z = 4000 + 600 * 2.05 = 5{,}230$

The reorder point should be the safety.

$R = Q^* - Q = 5{,}302 - 3000 = 2{,}230$

With this ordering policy, the safety stock is

$SS = \sigma z = Q^* - \mu = 5230 - 4000 = 1{,}230$, same as before.

2. The inventory level is three weeks ago. The current inventory level is 1100. How much should you order?

1100 is below the safety stock of 1230 by 130. This can be due to a large number of orders placed after last inventory review, or a single large order. Using this policy, you will still order 3000.

The expected cycle stock, inventory-related costs in this policy are similar to the fixed-cycle policy. In the fixed order cycles, the order quantity varies based on the current stock levels. In the fixed order quantity policy, the cycle times vary based on the demand. These two single decision policies, plus *EOQ*, illustrate some of the basic concepts in supply-chain engineering, and can go a long way to get some basic understanding among many inventory policies. These simple models help you to achieve solutions not too far from much more complex models.

2.4 Inventory turnover ratio

Inventory turnovers and the inventory turnover ratio are important concept in a supply chain.

$$ITR = \frac{Cost\ of\ goods\ sold}{Cost\ of\ average\ Inventory}$$

2.5. The strong impact of lead time uncertainty

The duration of a window of uncertainty itself can be uncertain. In the United States and most of the developed economies, the lead times in most cases have low uncertainty. However, certain industries are known to have high uncertainty. Construction delays are common due to many reasons. Construction uncertainty causes uncertainty in its upstream and downstream activities. In the developing economies, on the other hand, the lead time uncertainty is often high due to the uncertainties in supply, power availability, production system reliability, road conditions, and many other possible sources of uncertainty. The supply-chain professionals should understand that lead time uncertainty greatly increases the overall supply-chain uncertainty.

Definitions

Let us define three random variables.

1. The demand in unit time: D with distribution $D \sim D(\mu, \sigma^2)$. For example, the weekly demand follows $N(1000, 300^2)$.
2. The window of uncertainty: W with distribution $W \sim D(\mu_W, \sigma_W^2)$.
3. The demand during the window of uncertainty $D_W \sim D(\mu_{DW}, \sigma_{DW}^2)$.
4. D and W are independent.

Then, the standard deviation of the demand during the window of uncertainty is σ_{DW} is

$$\sigma_{DW} = \sqrt{\mu_W \sigma^2 + \mu^2 \sigma_W^2}$$

Example

A firm supplies a variety of tools. The demand for one of the tools follows $N(1000, 300^2)$ per month. There are two potential suppliers, Tortoise and Hare. Tortoise guarantees two months delivery, with no variability. Hare has an expected lead time of only 1 month but with a standard deviation of 0.6.

1. What is the standard deviation of the demand during Tortoise's lead time?
2. What is the same for Hare?
3. What is your conclusion?

Solutions

1. Tortoise l lead time is 2 weeks, constant.

$$\sigma_{DW} = \sqrt{\mu_W \sigma^2 + \mu^2 \sigma_W^2} = \sigma_L = \sqrt{2 * 300^2 + 1000^2 * 0} = 424$$

2. Supplier 2 demand and lead time are uncertain.

$$\sigma_{DW} = \sqrt{\mu_W \sigma^2 + \mu^2 \sigma_W^2} = \sqrt{1 * 300^2 + 1000^2 * 0.6^2} = \sqrt{90{,}000 + 360{,}000} = 677$$

3. The demand variability during lead time for Hare is 59% higher than that from Tortoise! Therefore, the lead time variability is very damaging. For that reason, the low-cost suppliers offshore must offer additional benefits to offset the lead time uncertainties.

Summary: Lead-time variability led to higher demand variability during the lead time than the demand variability. Conclusion: the variation in lead time led to higher cost than the demand variability.

Table 9.2. Variability levels in different situations.

LT \ Demand	$\boldsymbol{CV_D = 0}$	$\boldsymbol{CV_D > 0}$
0	0	0
Constant	0	H
$CV_{LT} > 0$	Higher	Highest

3. UNCERTAINTY IDENTIFICATION AND REDUCTION AND HEDGING

Today, firms often have plenty of historical data. Some even have access to people's digital footprint data. A lot of useful knowledge is embedded in the complex maze of data. People use data mining, data analytics. and data science to find actionable knowledge from these data. A lot of this effort is to reduce the uncertainty in decision-making. Although the data may appear uncertain, do not take the apparent variations for granted. As a supply-chain professional, you should work to identify underlining pattern or systemic factors behind the apparent uncertainty.

3.1 Reduce apparent uncertainty through analytics

Data may appear random. However, seemingly random data may have embedded systematic or deterministic variations. For example, what is the pattern in the figure below? The 72 sample numbers have the following parameter: Avg = 4.77, Std = 2.91, Min = 0, Max = 9. The distribution is shown in Figure 9.6.

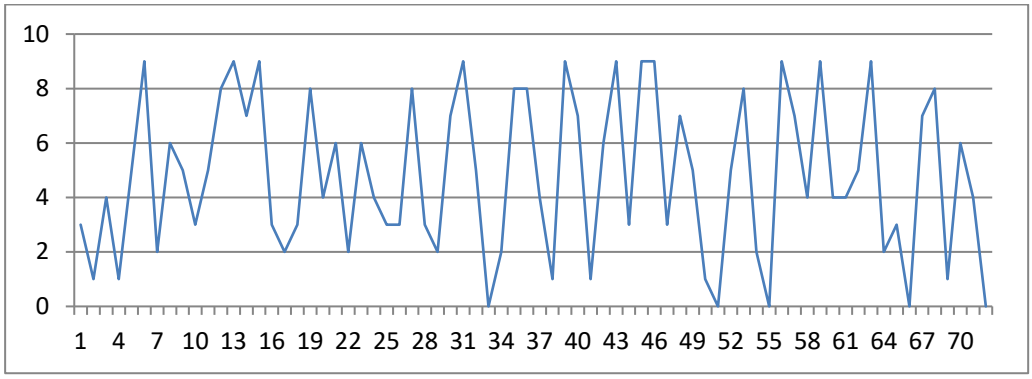


Figure 9.6. The value of subsequent digits of the value of π.

This series is the first 72 digits of π! Totally deterministic. In the first 6 billion digits, each digit appears about 10% of the time, or approximately follows a discrete uniform distribution. It possesses many characteristics of randomness, but it is deterministic!

Figure 9.7 shows a supply chain example of 122 days of demand of a paper product. Do you see any patterns? Closer observation can reveal that there is a weekly trend coupled with a monthly trend.

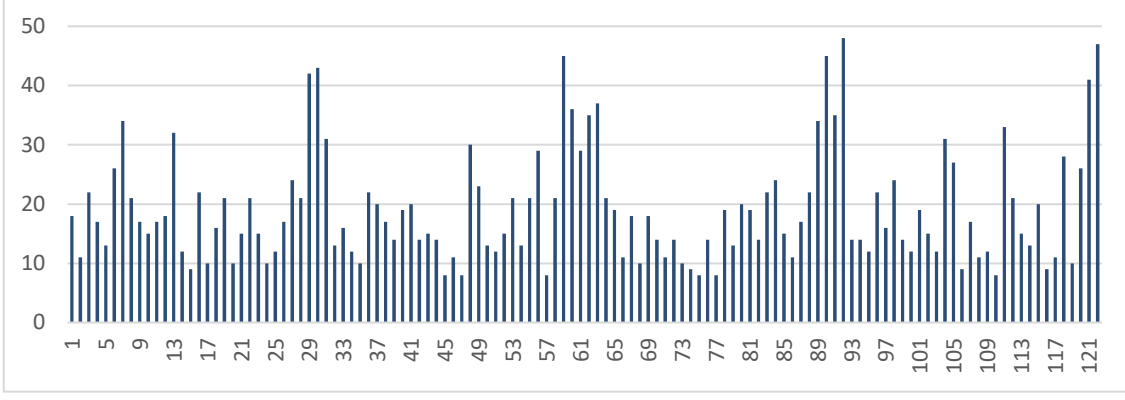


Figure 9.7. The demand of a product.

Example

A store is open Monday through Saturday. The demand for certain types of groceries in the last four weeks ranged from 0 to 6, with the specifics below. Please find

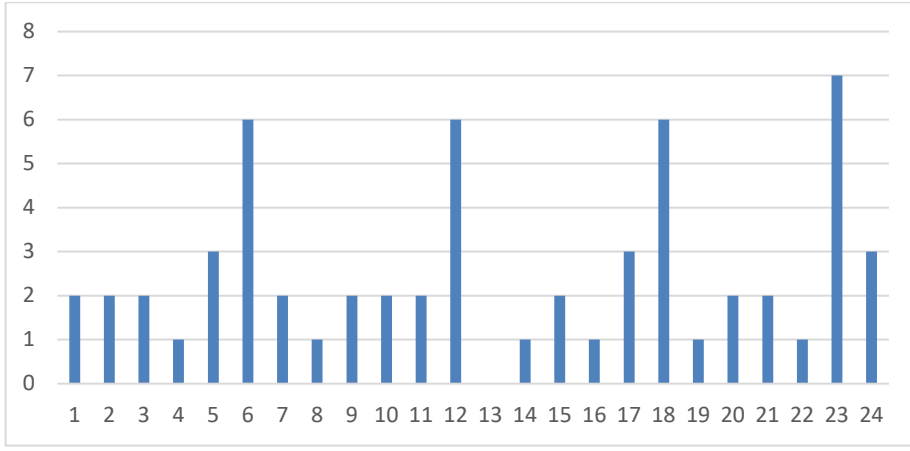
Days	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
D	2	2	2	1	3	6	2	1	2	2	2	6	0	1	2	1	3	6	1	2	2	1	7	3

1. The coefficient of variation of the daily demand.
2. Plot and identify the pattern.
3. Find systematic or deterministic variability and find the new coefficient of variation.

Solution

Assuming the demands are iid.

1. The average demand is $\mu = 2.5$. The standard deviation for the sample is $\sigma = 1.87$. Therefore, $CV = \frac{1.87}{2.5} = 0.75$, quite high for demand. Normal distribution may not be suitable.
2. The plot is shown below. It appears there is a weekly pattern. Friday and Saturday demands are higher. This is quite common due to people's work/break weekly patterns.



3. We can separate the weekday demand from the weekend demand. For weekdays, $\mu_{weekday} = 1.5$ $\sigma_{weekday} = 0.63, CV_{weekday} = 0.42$, $\mu_{weekend} = 4.5$, $\sigma_{weekend} = 1.93$, $CV_{weekend} = 0.43$. Both CVs are much lower than 0.75! The reason is you have removed the systemic or deterministic variation between the weekday and weekend from the raw data.

3.2 Reduce uncertainty in human processes

In the sharing section, we noted that human beings are very flexible. However, human flexibility also brings in uncertainty in the abilities in strength, height, cognitive capabilities, training, experience, and position on the learning curve. These will lead to uncertainty in performance in time, quality, and interruptions. The first thing is to identify the source of uncertainties. If the difference is due to strength, design the work contents to take strength out of the equation; and similarly in the other dimensions. You can also use job assignments to minimize the uncertainties and even reduce the overall uncertainty.

If the variation is due to different training, investing in training can be the solution. Variation can be due to interruptions, such as unplanned breaks, answering phone calls, and being called upon to help others. In a flow line, the alternate stoppage of stations can lead to many stops at different times. In this case, planned breaks can be used to reduce the uncertainty. If the uncertainties are due to the incentives, you may want to redesign the measurement and reward system discussed in other sections.

3.3 Reduce uncertainty in equipment

There are different aspects of equipment uncertainty: in process setup, changeover, shutdown, during operations, and due to interruptions. The setup, changeover, and shutdown processes are often manual and require skills and experiences. The methods of reduction are similar to those with manual processes.

The uncertainties associated with equipment processes can be related to the properties of the equipment themselves. Some are more predictable than others. Managers often focus on averages. However, the uncertainty in equipment performance can increase the cost tremendously. It may be worthwhile to invest in more predictable equipment.

The last factor is interruptions. Equipment requires input, needs space for output, may need operators to attend, may need a computer to control, and needs to be in the operating condition to function. If any of these are interrupted, the equipment cannot function. When the input material is fed wrong, the output is blocked; the operator, computer, or the equipment itself do not act properly; and the equipment operation will be interrupted. The interruption can be costly, especially in certain flow lines with expensive restarts, such as carpet, paper, cereals, and glasses. An interruption in one station will stop the entire line. Planning, training, process design, and preventative maintenance are all extremely important in such systems, because they can be planned, and can be done during non-peak times and off-line. It can be shown that regularly scheduled preventative maintenance at fixed intervals causes less productivity loss than unexpected interruptions (Hopp and Spearman, 2011).

3.4 Reduce uncertainty in input

In a production system, running out of supply can disrupt production. The Covid-19 pandemic led to a lot of such problems. Building relationships with the suppliers can go a long way. Most suppliers supply to multiple clients. At crunch time, they may favor certain clients over others. For unreliable supplies, or supplies from long chains, keep sufficient safety stocks.

To increase robustness, develop alternative suppliers for critical supplies. In long chains with many suppliers, watch out for the situation when the supplier's supplier may be the same.

In service, reducing arrival-rate uncertainty helps with planning the capacity and improving customer experience. On highways, the control of entrance to a steady rate helps to reduce congestion. In restaurants or medical services, reservations or appointments accomplish the same.

3.5 Buffer demand uncertainty with time, resource, or capacity

In most situations, demand is uncertain. Buffering can be a good strategy to deal with such uncertainties. For example, the generic screw on a laptop can be buffered by inventories. Such

screws are low in cost and long in supply chain lead time, and with potential variation in supply lead times.

3.5.1 Buffering with inventory

For the supplies to steady and long-term demand, with low cost, long lead times, and significant quantity discount, it may be most prudent to buffer with inventories, such as screws.

3.5.2 Buffering with time

In many products or services, inventory may not apply. For example, you cannot stock up repair, services. If the supply is expensive, can be acquired quickly, or changes over time, carrying inventory is expensive. In such cases, a good strategy can be buffering with time. It means to ask the client to wait a longer time. The conditions may be true for any supplier. Therefore, the acceptable lead time for the customer is long anyway.

3.5.3 Buffering with capacity

Some products or services cannot wait, such as emergency services, including equipment, personnel, and supply. In a fire department, the demand cannot be buffered or wait for service. The only way to buffer is through capacity or sharing (in the last chapter) or pooling (in the next section).

3.6 Artificial scarcity to speed up the demand

An out-of-box strategy to deal with demand uncertainty is to limit the supply by design. This is called artificial scarcity. This requires the supply to be sufficiently differentiated from the competition. The first company that successfully adopted this strategy in large scale was Nintendo. When Nintendo developed the Game Boy games in the 1990s, it was greeted with overwhelming success. Nintendo soon could not keep up with the supply. In this situation, most companies would first expand the supply capacity to meet the increasing demand. Nintendo, instead, took a measured expansion. They kept the supply slightly less than the demand. As a result, Nintendo can produce without considering the demand uncertainty. This way, there is no need for safety stock. Since the competition is so different, and the game market enjoys the network effect, they do not have to worry about lost sales. Ultimately, they can realize the total sales as if there were no artificial scarcity. In the process, they reduce cost in production through planning, eliminate safety stock, avoid potential obsolescence, and the biggest impact is that they can keep the prices high. Nintendo repeated the same feat with Wii in the 2000s.

Costco adopts artificial scarcity strategy for its seasonal or flash sales of products in fashion, electronics, recreational devices, furniture, and so on. Costco intentionally orders less than expected demand to create artificial scarcity. The scarce supply limits total sales to the demand and avoids excess inventories. In addition, customers know when they see something they like, they must buy it right away. The product may run out before the next visit. This also reduces the choice overload and speeds up the sales. For a retailer, the important measure is the total revenue received, not the revenue from one single item. The faster sales allow the display for the next product. As a result, Costco can create more revenue per unit time per square foot. The space efficiency is related to real estate, utilities, and personnel. If you measure the sales per square foot, Costco is by far the highest among discount (warehouse) stores: Costco \$1,100/ft^2, Sam's

Club $680, and Walmart $400, only above one third of Costco's (from Huffington Post, May 29, 2014, http://www.huffingtonpost.com/2014/05/29/costco-earnings_n_5412588.html).

Online retailer Rue Lala setup Rues sales with window times. They display their price, and a reference price, and numbers of item left. When you see something you like, you must buy soon. Otherwise, Rue Lala will take the item out after some time. Rue Lala asks you to log in to browse, and following your browsing and shopping habits, uses them to determine what to show to you. They use this artificial scarcity to increase sales rate.

4. UNCERTAINTY REDUCTION THROUGH POOLING

Pooling is related to sharing discussed in Chapter 8. We have already seen that the relative uncertainty of book sales in a year is lower than in a week. You may also notice that the number of emails you receive each hour can vary a lot. However, the variation may be less in a day, and even less over a week, and so on. Also, the relative uncertainty in the sales of bottle of soda in a plant is lower than on a vending machine. These are due to the pooling effect.

4.1 Mathematical basis for the pooling effect

Consider that there are n random variables, $X_1, X_2, \ldots X_i, \ldots X_n$, with their corresponding mean μ_i and standard deviation σ_i. If we pool these random variables together, or sum them up, the mean, standard deviation, and CV for the pooled variables $X = \sum_i X_i$ is

$$\mu_{pool} = \sum_i \mu_i$$

$$\sigma_{pool} = \sqrt{\sum\nolimits_{i,j} COV_{i,j}} = \sqrt{\sum\nolimits_i \sigma_i^2 + \sum\nolimits_{i \neq j} COV_{i,j}}$$

$$CV_{pool} = \frac{\sqrt{\sum_i \sigma^2 + \sum_{i \neq j} COV_{i,j}}}{\sum_i \mu_i}$$

If Xs are not correlated, the COV between different variables are all zeros, then

$$CV_{pool}^{no\ corr} = \frac{\sqrt{\sum_i \sigma^2}}{\sum_i \mu_i}$$

Figure 9.8 shows the trend of means, standard deviation, and CV for the sales of an item. with daily sales of one ton and daily standard deviation of 3, assuming the daily sales are not correlated. The expected values grow linearly. The standard deviation grows with the square root of number

of says, and CV decreases from 3 (very high variability) to 0.5 (low variability) from a single day to a month.

What is the mathematical basis for this drastic reduction? The expected value grows linearly while standard deviation grows at the rate of square root, shown in Figure 9.8.

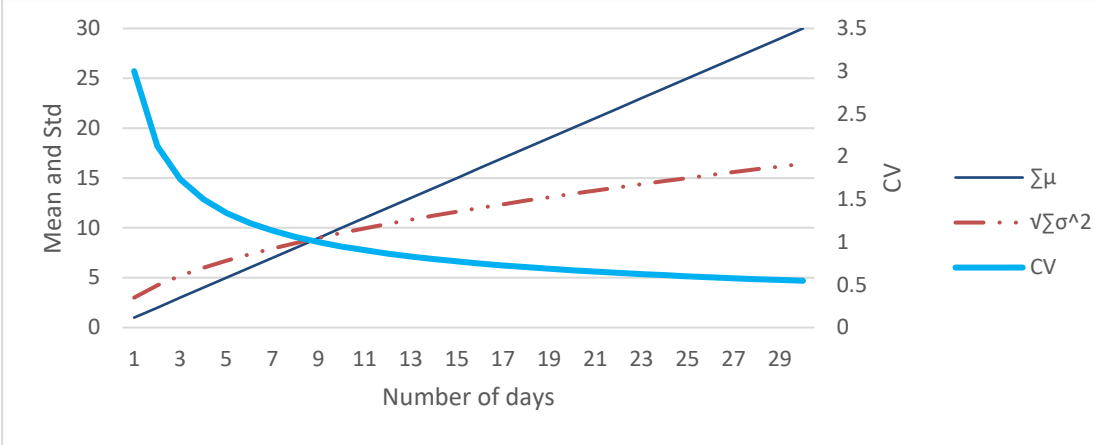


Figure 9.8. The change of mean, standard deviation, and CV of demand with longer time intervals.

4.2 Special case of pooling among identical and uncorrelated random variables

When each of the *n* random variables are:

1. Independent and identically distributed with μ, σ^2, such as the sales on a vending machine in a day.
2. Not correlated with each other, such as the demand in multiple days or from multiple vending machines.

Then we have

$$CV_{pool}^{identical} = \frac{\sqrt{n\sigma^2}}{n\mu} = \frac{1}{\sqrt{n}}\frac{\sigma}{\mu} = \frac{1}{\sqrt{n}}CV$$

This means that when *n* identical random variables with same mean and variance are pooled or combined, the coefficient of variation is reduced by a factor of $\frac{1}{\sqrt{n}}$. In a supply chain, this can be a good approximation when pooling among multiple time periods, stores, customers, and so forth.

4.3 The Impact of Correlations

In practice, the daily sales can be correlated. If the correlations are known, you can use the more complex form. Note that the correlations can be both positive and negative. People who buy today may be less likely to buy tomorrow, or even in the near future. However, a single person's

purchase may lead to others who also purchase. Therefore, correlations can be positive, negative or 0. You can show

$$0 < CV_{pool} < \max_i\{CV_i\}$$

The $CV_{pool} \to \max_i\{CV_i\}$ can occur only when the random variables are perfectly and positively correlated, which is no longer random.

If two stores of the same company carrying the same products, the sales in both stores will go up when the company runs a promotion. In this case, $COV > 0$. The pooling effect will be reduced. You can see in the example that even when that is the case, the pooling effect can still quite strong.

If the two stores are independent, the one runs the promotion and the other not. One will experience increased demand while the other may experience reduced demand. In this case, the $COV < 0$, and the pooling effect is even stronger. If two random variables with the same mean are perfectly negatively correlated, $CV = 0$. However, the process is no longer random.

Therefore, quantitative modeling must be associated with proper assumptions. Simplistic generalizations can be dangerous. Readers are encouraged to calculate the coefficient of variation of two stores with perfect positive correlation and negative correlation.

4.4 Dimensions of pooling

Pooling can be used to reduce uncertainty in many different dimensions. We have already seen pooling over time, from daily sales to monthly sale in a vending machine, or pooling among supply channels from a bottling plant. Here is a simple summary of the dimensions of pooling. Many of these have already been discussed in the prior chapter on sharing. In sharing, the focus was the benefit of using the averages. The focus of pooling is the reduction in relative variability.

1. **Time**: The longer the time duration, the lower the variability or uncertainty. Please consider the demand of a coffee shop on a daily, weekly, monthly, or annual basis. Do you know how to quantify?
2. **Space**: Pooling can also occur in space. The variability of space requirement in randomized storage is much lower than in dedicated storage due to the pooling effect.
3. **Channels**: Pooling can be applied to one SKU when channels are combined, as in the above example.
4. **Material handling or transportation**: SKUs can be pooled to reduce the variability in handling of pallets, trucks, and so on. You can have mixed truckloads, mixed pallets, or mixed cases that enjoy the pooling.
5. **Machines or workstations**: A pool of identical machines with variable randomness exhibits less variability as a whole.
6. **Personnel:** You can have a cross-trained workforce to pool and reduce the problem with demand variability.
7. **Among countries with different currency:** More later when offshoring is discussed.
8. **Among suppliers:** Firms that source from different suppliers can enjoy the pooling effect of the risks.

4.5 Pooling examples

4.5.1 Pooling in time

We have already seen pooling in time in the newsvendor problems. Given the weekly demand $D(\mu, \sigma^2)$, the demand during the window of opportunity W, if the weekly demands are uncorrelated, is $D(W\mu, W\sigma^2)$. The coefficient of variation is reduced by $\frac{1}{\sqrt{W}}$.

Example

Cars have many options. For example, a 2020 Chevrolet Malibu has LS, RS, LT, and Premier models, and each has many colors and other options. One of major supply strategy decisions is the promised time to complete the orders. Shorter times provide the customer with a better experience. However, a longer time provides the seller longer time to plan. Let us quantify the benefits of longer time. Consider a particular model. The order rate is one car per hour by a customer somewhere through a dealership or online. Consider that there are 9 hours per workday, and 6 days per week. Please find the demand variability during the following planning horizons.

1. 1 day.
2. 1 week. This is useful, because many manufacturers deliver once a week to dealers.
3. 8 weeks. This is an actual strategy by a supplier.

Assumptions

We can assume the order arrival times are memory-less. Exponential distribution would apply.

1. The daily arrival rate follows $Pois(\lambda t) = Pois(9)$. The $CV_{day} = \frac{1}{\sqrt{9}} = 0.33$.
2. Another useful measure is weekly. The deliveries to different regions are often done weekly. $CV_{week} = \frac{1}{\sqrt{9*6}} = 0.136$.
3. The supplier promised 8-week delivery. $CV_{week} = \frac{1}{\sqrt{9*6*8}} = 0.048$, near constant.

In reality,

- The daily demand may not be iid. Someone who buys today would not buy tomorrow. The demand in successive days can be negatively correlated due to special situations such as a holiday or bad weather.
- The demand within a week tends to be seasonal.
- The salesperson has monthly quotas. As a result, there may also be monthly patterns.

4.5.2 Pooling in space

You can pool the utilization in space. One of the major costs in college education is classrooms. Twenty years ago, Georgia Tech had more, maybe less polished, classroom spaces. The university allocates classrooms to the departments. There were plenty of unused classrooms then. As the enrollment and need for quality grows, the classrooms became the bottleneck. Georgia Tech applied space planning to allocate all classroom spaces on campus. The utilization, or the effective capacity, increased tremendously.

One of the major costs in a warehousing operation is the storage. In warehousing, you can adopt a dedicated storage strategy. It provides convenience in searching. However, in dedicated storage, an empty space dedicated for one SKU cannot be used for another, and this leads to wasted space.

You can employ randomized storage to enjoy the pooling effect. An empty space can be used for any SKU. So, the space requirements of different SKUs are "pooled." To enjoy the pooling effect, one must have a way to track the location of the SKUs. Warehouses with a Warehouse Management Systems (WMS) often employ randomized storage for the reserve, or low-velocity products, and a hybrid strategy for high-velocity SKUs. Retailers normally use a dedicated method to display merchandize using dedicated storage strategy.

Example

A distribution center handles pallets. It has 1,000 SKUs in the class B category. The average inventory level is 10. The actual inventory level can be approximated by uniform distribution $U(0, 20)$.

1. If the company would allocate 95% space for each SKU, how many pallet locations are needed in storage? This means a pallet of a specific SKU has a 5% probability of being without dedicated locations, although there may be dedicated storage of a different SKU.
2. If the company adopts randomized storage, how many pallet locations are needed?

Solutions

1. For dedicated storage. For each SKU, you will need 0.95 times the maximum needs. You can then multiply by all SKUs, or

$$20 * 0.95 * 1{,}000 = 19{,}000$$

2. For randomized storage, a space is a space. We assume that the replenishment and demands of different SKUs are independent. The sum of 1000 uniform distributions can be approximated by normal distribution. We can find the parameters of the normal distribution as follows.

The space requirement for randomized storage is

$$\mu = 10, VAR = \frac{1}{12}(b^2 - a^2) = \frac{1}{12}(20^2) = 33.333, \sigma = 5.77$$

$$\sigma_{total} = \sqrt{1000}\sigma = 31.62\sqrt{\frac{1}{12}(b^2 - a^2)} = 31.62\left(\sqrt{\frac{20^2}{12}}\right) = 182.56$$

$$\mu + \sigma z = 10{,}000 + 183\Phi^{-1}(.95) = 10{,}000 + 183 * 1.64 = 10{,}300.$$

This represents 84% space savings to let the pallet share 100%. Even if you increase the percentage of space to 99.9%, the total space requirement is only

$$\mu + \sigma z = 10{,}000 + 183\Phi^{-1}(.999) = 10{,}000 + 183 * 2.99 = 10{,}547$$

That is still an 80% savings!

This example shows the idealized situation. In practice, there are safety stocks that reduce some savings, and the replenishment and demand may not be independent and may also need more space. However, the space utilization can reach 85% or 90%. Therefore, there is significant space savings in randomized storage.

4.5.3 Pooling among distribution channels

We can pool among supplier or distribution channels.

Example

A distribution center supplies to 100 retail stores. The weekly demands for bottled water at all stores have the same average of 100 cases with a variance of 60^2. The demands from different stores are independent.

1. What is the coefficient of variation for each store?
2. What is the coefficient of variation for the DC if the demands from stores are independent, or their $COV = 0$?
3. If the demands among the stores are positively correlated, what should be the range of the coefficient of variation for the DC?
4. If the stores are in pairs, and the demands between pairs are independent but the demands between the two stores in the pair are negatively correlated, what should be the range of coefficient of variation?

Solutions

1. The coefficient of variation at a store is

$$CV_{Store} = \frac{60}{100} = 0.6$$

2. The coefficient of variation at the distribution center is

$$CV_{DC} = \frac{\sqrt{100*60^2}}{100*100} = \frac{\sqrt{100}*60}{100*100} = \frac{1}{\sqrt{100}}\frac{60}{100} = \frac{1}{10}0.6 = \frac{1}{10}CV_{Store}$$

3. If the demands are positively correlated, the CV for the DC should be higher than calculated above. The maximum will be CV for the store. It will require that all stores vary in unison, which is almost impossible.
4. If demands are negatively correlated, the CV for the DC will be lower than calculated above. If each pair is perfectly correlated with a coefficient of -1, the CV for the DC will be zero.

4.5.2 Pooling among correlated channels

In practice, the demands are often correlated. If the demands are correlated, you should first find the covariance. With the consideration of correlation, the data set tends to be very large. Let us use a very small data set to show the correlated demand. The data is not statistically significant.

Example

A distributor supplies to three stores. The demand in the six days in last week were

	D_1	D_2	D_3
1	13	10	4
2	1	1	18
3	4	1	11
4	10	11	9
5	7	6	6
6	11	12	2

Please find

1. The coefficient of variation for each store.
2. The coefficient of variation if we assume that the demand among the three stores is independent.
3. The coefficient of variation considering the correlation.

Solutions

1. $VAR_i = E[X^2] - (\mu_i)^2$. You will find the result is smaller than what Excel calculates. This is because of biased estimate versus unbiased sample.

μ	7.67	6.83	8.33
VAR	20.67	24.57	33.07
CV	0.59	0.73	0.69

2. If we assume that the three demands are not correlated, the coefficient of variation will be the square root of the sum of the variances divided by the sum of the means, or

$$CV = \frac{\sqrt{20.67+24.57+33.07}}{7.67+6.83+8.33} = 0.39$$

This is smaller than any of the three CVs individually. This is the pooling effect.

3. $COV(X,Y) = E(XY) - \mu_x\mu_y$. For D_1 and D_2, you multiply the demand for each day, then find the average, and subtract the product of means, which will give you 69.83 -7.67 * 6.83

= 17.44. You can find the pair between 1 & 2, 1 & 3, 2 & 3, 2 & 1, 3 & 1, and 3 & 2, or you can put everything in a matrix.

	D_1	D_2	D_3
D_1	20.67	17.44	-19.22
D_2	17.44	24.57	-18.94
D_3	-19.22	-18.94	33.07

This is a symmetric matrix. The cell with D_1 and D_1, is VAR_{11}. The cell with D_1 and D_2, is COV_{12}. It is clear that D_1 and D_2 are positively and strongly correlated, while D_1 and D_3 are negatively and strongly correlated. D_2 and D_3 are also negatively and strongly correlated, which should make sense. The numerator for CV will be the sum of the nine elements, while the denominator is the sum of the means, or $CV_{Pooled} = 0.27 < 0.59$. This means overall, the demands are negatively correlated.

4.6 Pooling effect and bullwhip effect

Have you heard the term "bullwhip effect?" It refers to the fact that as you move upstream in the supply chain, the variability becomes higher and higher. One can experience this phenomenon by playing a supply chain game, sometimes called beer game in-person or through virtual environment through some applications.

The pooling effect in distribution above suggests that the variability becomes lower and lower as you move up the supply chain due to pooling. This violates the bullwhip effect. Where is the inconsistency?

Actually, the bullwhip effect occurs in a chain, between one supplier and one client, shown in Figure 9.9. The variable lines indicate the demand variation over time at the store and at the DC. The bullwhip effect says that the variation at the DC is higher than at the DC, expressed in higher variability in the height of the line in the DC. The average should be the same based on the conservation of mass. In addition, the demand hits the store at a higher frequency, probably in minutes, than the orders the store sends to the DC, which may be daily, weekly, or even at longer time intervals.



Figure 9.9. Bullwhip effect from one store to the distribution center.

Many reasons cause the bullwhip effect.

1. The demand frequencies at downstream is high with small quantities in each demand. However, the orders that downstream makes to the upstream are lumpy and less frequent.
2. Both the lead times and order cycle cause delay in signals between downstream and upstream.
3. Human psychology and overreaction to shortages and overstocks.

The bullwhip effect increases with more firms along the chain. In the beer game there are a producer, a distributor, a wholesaler and a store. The effect can also be enlarged by large demand shocks at the end of the supply chain.

However, the pooling effect occurs in a supply tree, shown in Figure 9.10. A DC typically supply to many stores, and a factory supply to many DCs. The combined channels or lanes add random variables together. Both the mean and standard deviation go higher, but the mean increases faster than the standard deviation. The result is higher in average but lower in CV, as shown in Figure 9.11. This effect can also be enlarged with multiple levels of pooling. You can refer back to the soda sales from a vending machine versus from the bottling plant.

If the pooling is from many channels, the pooling effect is most likely stronger than the bullwhip effect. In later sections, we will also look at other strategies to reduce the bullwhip effect.

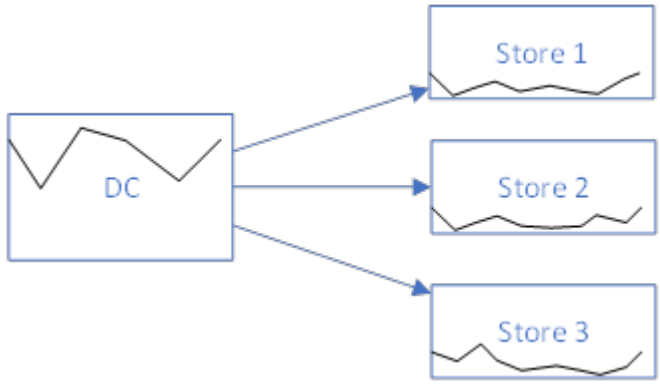


Figure 9.10. Bullwhip plus pooling effect.

5. VARIABILITY AND UNCERTAINTY IN BARILLA CASE

Please refer to the Barilla case study. The focus of the case study is that the Barilla's existing strategies increase the variability and uncertainty in the supply chain. The distributor orders the pastas from Barilla and sell them to the independent supermarkets. The coefficient of variation of weekly sales is low at approximately 0.2. Note, this low *CV* is pooling over 100 Barilla SKUs. The *CV*s for individually should be higher. The DC carries 2.7 weeks of inventory. However, it

still experiences significant stockout between 5 – 9 percent even they order every week. However, *CV* of the weekly orders from the DC to Barilla increased to over 0.75. This is bullwhip effect. Barilla can change it pricing, information and incentive strategies to reduce the uncertainty. A smoother demand from the DCs can help to reduce the setup cost in production and transportation cost in supply chain are high. The new pricing strategy can also help to reduce the prices at the supermarket to increase Barilla's competitiveness.

6. SUMMARY OF UNCERTAINTY

Uncertainty is part of the supply chain. The major sources of uncertainty can be in demand, human behavior, partner behavior, and process interruptions. To be competitive, a supply-chain professional should find ways to quantify the level of uncertainty, model the uncertainty, and to find strategies to reduce or cope with the uncertainty through collaborations, incentives, pooling, buffering, and so on. The coefficient of variation is an excellent and simple way to communicate the level of uncertainty. The newsvendor model provides a simple way to think about the cost of uncertainty and the solution strategies. It actually provides good solutions for many practical problems.

Several factors that paralyze the ability to deal with supply-chain uncertainties are the innate variability itself, reflected by σ; the setup time and costs, reflected by K; and the window of uncertainty, W. A supply-chain professional should consider process design, training, incentives, collaborations, and so forth to reduce σ; to invest to reduce K; and to reduce the window of uncertainties, W.

7. THE BENEFIT AND COST OF SUPPLY-CHAIN UNCERTAINTY TO HUMAN NEEDS

7.1 The benefits of uncertainty

Uncertainty provides business opportunities for those who can deal it well or can help others to deal with it, such as in the insurance and financial industries. It also provides job opportunities for engineers. One of the most important objectives of analytics or data science is to identify predictable behaviors from complex information from diverse set of sources now become more and more available through smart devices, sensors, vision systems, the internet of things (IoT), and historical records and plans. These increase employment or livelihood for humans.

For adventurous people, some level of uncertainty provides a thrill to enrich their lives.

7.2 The cost of uncertainty to human needs

Safety is part of basic human needs. Uncertainty can reduce the level of safety. Based on the simple newsvendor model, uncertainty always adds cost to the supply chain. To avoid stockouts against uncertainty, the supplier adds safety stocks at extra cost, or to plan for extra capacity in supply with safety resources.

The supply-chain cost can be very high for the rare unpredicted events such as great recession of 2008, the Covid-19 pandemic, natural disasters, or the ocean liner that got stuck in the Suez Canal. Even for those who can benefit from it financially, the excessive added stress can lead to various challenges to their basic human needs such as stress and depression.

10. ECONOMIES OF SPEED AND NETWORK DESIGN

"Time is money" is literally true in the supply chain. If you can differentiate by supplying faster, you can charge more to increase revenue. Furthermore, if you can supply anything instantaneously, there is no need to worry about uncertainty in demand or supply. There is no need for buffers, inventories, plans, and so forth. However, no one can supply physical goods that fast. So, supply lead time cannot be eliminated. However, faster speed can bring higher revenue or reduction in stocks and resources. The supply-chain professionals should always consider speed as part of supply-chain strategy. However, when you communicate your ideas of shorter lead time and faster service to upper-level management, you should learn to convert shorter lead times or faster services into financial terms.

1. THE BENEFITS OF FASTER SUPPLY

The internet and lean drive have speed up the supply of products and services. Faster supply has many direct benefits.

1.1 Improve client's supply chain

If you can supply faster than your competitors, you can provide perceivable benefits to your clients. For a spot market, the faster supply may allow the client to accomplish their task faster. For the long-term recurring supply of make-to-order supplies, the client can enjoy a shorter window of uncertainty, which has been shown to lead to lower cost due to uncertainty. Remember that the window of uncertainty is the sum of order lead times and order-cycle times. The faster supply can reduce the client's leadtime, safety stock and stockouts and differentiate from the competitors.

1.2 Increase prices

Faster supply can possibly be worth a higher value to the customers. As a consumer, many people are willing pay more to get faster deliveries, directly or through memberships.

2. FORECAST ACCURACY INCREASES WITH SHORTER FORECASTING HORIZONS

Firms must forecast demand and capacity for their planning. The shorter the forecasting horizon, the higher the forecasting accuracy because of additional information that can be used to forecast. In uncertainty, we have assumed that the demands in different intervals during the planning horizon l are iid. That means $\sigma_1 = \sigma_2 = \cdots = \sigma_l$, shown in Figure 10.1.



Figure 10.1.Illustration that forecast errors are iid over the forecasting horizon.

2.1 Inaccuracies in pooling in time

Based on the iid assumption, the demand variability during the entire planning horizon can be pooled. For example, if the daily demand forecast error is σ, then the demand forecast error for five days is

$$\sigma_5 = \sqrt{5\sigma^2} = 2.24\sigma$$

2.2 There is more information in shorter forecasting horizons

In practice, more information is available for tomorrow than the days after tomorrow because we can already learn what happened in the market and in the supply chain today, and can use this to improve the forecast for tomorrow. Therefore, the forecast errors increase over time, as shown in Figure 10.2.The forecasting error increases faster than the results from iid and the pooling effect.



Figure 10.2. Forecast errors increase with the forecasting horizon.

You may be able to link this rationale with weather forecasts offered hourly, daily, five-day, ten-day, and monthly. The longer the forecast horizon, the lower the accuracy. Smith and Sincich (1991) studied the effect of length of forecast horizon. Their empirical results show that the forecast accuracy decreases almost linearly with the increase of forecasting horizon. If we apply this to the case above,

$$\sigma_W \approx 5\sigma \gg 2.24\sigma$$

Therefore, the pooling effect estimates for the window of uncertainty is an underestimate. This should be considered in the estimation of benefits of faster services.

3. THE COMPONENTS OF ORDER FULFILLMENT TIME

We define the order fulfillment time as the time from when an order is received by the supplier to the time the order is delivered to the client. The supplier must ensure that order fulfillment time must be less than or equal to the customer's acceptable lead times, or

$$LT_{Order} \leq LT_{Acceptable}$$

Order fulfillment time is important for the client to plan for their activities. The supplier can be a manufacturer or a distributor. There are many components that constitute the time needed for the supplier to deliver an order. We will discuss a few that are important for supply-chain design strategies.

3.1 Production time

The production time is from the time to set up the system to produce an order to the time the order is completed. The term "production" is generic. It can be the assembly of a car or a chair, it can be to pick up an order in a distribution center, or it can be to compete a haircut. A car takes about a day to assemble; a dealer's order of 10 cars can be completed in 2 days (the cars in one order may not be processed in succession); a customer order takes about an hour to be picked in a warehouse; or a haircut takes about 30 minutes.

Note that for a make-to-stock supplier, the direct production time perceived by the client is the order-picking time. The actual production time is internal to the producer. It impacts its own replenishment frequency and quantity.

The actual production time can be further divided into setup time and running time to complete the lot. In the economy of scale, we have discussed that the lot should be near EOQ to minimize the cost. The lot size, or the production time, increases with the square root of setup time and cost, or $EOQ \propto \sqrt{K}$ and $CT \propto \sqrt{K}$. If you can reduce the setup time and costs, you can reduce the economic order quantity and the product cycle time. This can reduce the internal cycle stock. It can also reduce the safety stock due to a shortened window of uncertainty; recall $W = LT + CT$.

Example

A company orders products from overseas. The demand is 200 units per month. The ordering leadtime is 1 month. The ordering and transportation cost is \$400 within the quantity range. The holding cost is \$1 per unit per month. The variance is 60^2 for each type.

1. What are the cycle time and lead time?
2. If the company can apply various kinds of automation to reduce the setup and fixed part of the transportation cost down to \$25 per setup, what should be the cycle time before and after?

Solution

Assumptions

EOQ assumptions approximately hold.

1. If setup cost is \$400,

$$CT = \sqrt{\frac{2*400}{1*200}} = 2 \text{ months}$$

2. If setup cost is \$25

$$CT = \sqrt{\frac{2*25}{1*200}} = 0.5 \text{ months}$$

The results are in months. It is clear that setup time and cost can be a major hurdle to speed, and it is often not obvious. The setup time and cost may only be a small fraction of the production time. However, their influence is high. What are the possible ways to reduce setup time and costs in the supply chain?

Computers and the internet can help to shorten the setup times and cost in order processing, production setup, order picking, and so on. For example, a long time ago, people who sent packages through UPS would fill out a form or input information to a handheld unit. This takes quite some time. Today, a user can setup an account, and keep most of the information unchanged. For each shipment, just enter the direct relevant information before the pickup or drop-off. There are two parts of this process that helps to reduce the setup cost. The first is to maintain the permanent information in the system. The second is to move most of the information process offline. This is true in many modern supply chain-order processing systems.

A further reduction in the repetitive orders is to let the status of the system trigger the orders automatically. For example, one person leaving the store would automatically send a message to the next person; or when the inventory level drops below the reorder point, order a pallet, or each Monday at 8 place an order to bring the inventory up to level.

3.2 Production planning time

When a make-to-order supplier receives an order, it will plan to produce the new order upon completion of current orders and the arrival of materials needed to produce the order after its order lead time. The more orders waiting in the queue, or the longer its material order lead time, the longer the new order has to wait before its production. The supplier would like to keep a queue to improve its facility utilization.

For a make-to-stock supplier that produces *n* types of products, it may produce them alternatively or cyclically. The cycle time to produce all product types is approximately the sum of the production cycle time for all product types, or

$$CT_{all} \sim \sum_{i=1}^{n} CT_i \propto \sum_{i=1}^{n} \sqrt{K_i}$$

You can see the importance of setup time and costs, discussed in the economies of scale.

Example

A supplier makes 30 types of casters. The EOQ dictates its cycle times fall into three categories, 1 day, 2 days, and 3 days, and each has 10 types.

1. If the firm builds each type cyclically, what is the cycle time to build all types of casters?
2. The firm has an opportunity to reduce the setup cost and time by 75%. How much influence is this to its cycle stock, safety stock, and total cycle time?

Assumptions

The information given is rather limited. We assume:

1. The information given captures the essence for the purpose of this simple analysis.
2. We also assume the EOQ and newsvendor model hold.
3. The casters only need commodity products, plastic pallets, steel bars, balls in the bearings, and so on. These are good candidates to buffer with inventories. Therefore, the cycle time is dominated by the production-cycle times.

Solutions

1. The approximate total cycle time is approximately

$$CT_{total} \approx 10(1+2+3) = 60$$

2. If the setup time/cost can be reduced by 75%, the new cycle time

$$CT_{total}^{New} \approx \sqrt{\frac{K^{New}}{K}}\, CT_{total} = \sqrt{\frac{0.25K}{K}}\, CT_{total} = 0.5CT_{total} = 30$$

3. The cycle stock will also be reduced in half.
4. The safety stock $SS{\sim}\sigma_W z$. $W \approx LT_{total} + CT_{total}$. Here, the leadtimes are all internal to the firm and are under control. It can be kept must lower than the cycle times. Then,

$$SS{\sim}\sigma_W z = \sigma\sqrt{CT_{total}}z$$

$$\begin{aligned} SS^{New}{\sim}\sigma_W^{New} z &= \sigma\sqrt{CT_{total}^{New}}z = \sigma\sqrt{0.5CT_{total}}z \\ &= 0.707\sigma\sqrt{CT_{total}}z \\ &= 0.707SS \end{aligned}$$

The reduction in setup times, or making the system more flexible, or increasing the economies of scope, have so much advantage just for the internal inventory alone. There are also other benefits.

3.3 Shipping time

Shipping time is from the time to load, transportation, and to the completion unload. The TL shipping time from Atlanta to Charlotte can be 6 hours, all inclusive.

One way to speed up the shipping time is by changing the shipping modes. For long distance with significant drayage, change from rail to truck can reduce the time from a couple of weeks to a couple of days. The total cost will increase but may be justifiable. The change from truck to air in long distance can reduce the shipping time from days to overnight. In this case, the cost increase is significant.

Apple ships their products from southern China to a location in United States by air. The shipping time goes down from many weeks to a day. However, the cost of air freight is much higher than ocean shipping. To mitigate the cost explosion, Apple ships the phones with minimum packaging and other unnecessary parts. The product will be "finished" in a U.S. facility before it is shipped to its destinations. In this way, Apple can achieve much faster speed with acceptable increase in shipping cost. However, the economies of speed improved supply chain in many other parts of the supply chain for their high value and high profit products.

3.4 Shipping planning time

Similar to production, there is also a shipping planning time. A TL load shipping time from Atlanta to Charlotte is 6 hours. However, the service may be once a day. Therefore, the planning time is one day minus the shipping time. A distributor has a contract with the TL service provider every week. The planning time will be one week minus 6 hours.

In the Barilla case, Barilla pays for the transportation cost from its CDC to the distributor's DC. To save the transportation cost, it offers a 2–3% discount. The salespeople can offer an additional 4% discount if a distributor orders three truckloads. Recall that Barilla runs promotional campaigns with sales targets. If a sales representative has a large gap from the target, he or she would push for three truckloads, on whatever in stock. As a result, the logistics manager at Barilla have to adjust is delivery plan. Therefore, the logistics manager needs longer and more flexible planning time to satisfy highly variable delivery demand and still experience high inventory and stockouts.

If Barilla offers frequent deliveries, such as less than a week with tighter delivery window, The impact to the inventory levels and stockout can be significant. Speed can play a very important role in the efficiency and service quality of a supply chain.

4. THE ECONOMIES OF SPEED FOR HUMAN NEEDS

At a macro-level, we will consider the impact of faster supply for satisfying human needs.

4.1 The benefits of speed

Faster service can get what humans needs faster. In emergency situations, faster supply can deliver air, water, food, shelter, and medicine faster to save lives and reduce anxiety and suffering. The faster supply reduces the customer wait times, reduces the storage requirements, and reduces the funds tied up in storage.

The faster speed at any link can also enable the faster speed at other links in a supply chain. Imagine a supplier who can change deliveries from every week to every day. This will enable the client to serve its customer faster.

4.2 The drawbacks of speed

However, ever-faster motion can potentially increase the severity of accidents. The faster pace in life adds anxiety to some. As society becomes more affluent, faster supply beyond human needs may not add as much value as what we lose in sleep and in time to think and to enjoy. The food from a faster dominant supply chain of certain food is lower in nutrition and taste, leading to more environmental degradation. Examples include chicken farms, CAFOs, and GMO seeds. Some of the negative impacts are not part of the financial accounting system, such as health, pollution to air and water, and degradation of land.

In March 2019, the FDA lifted an important restriction on genetically engineered salmon to allow AquaBounty, a biotech company with facilities in Canada and Panama, to start marketing the first GMO seafood. They grow twice as fast. However, anyone who has tried farm-raised salmon and wild salmon knows that faster growth negatively impacts flavor and nutritional value.

5. SUPPLY-CHAIN NETWORK DESIGN

In network design, we must consider speed as well as scale, and sharing and pooling effects to satisfy the client's need with minimum cost. We have used an illustration throughout. Here, let us reuse it to define the definitions show the relationships in Figure 10.3.

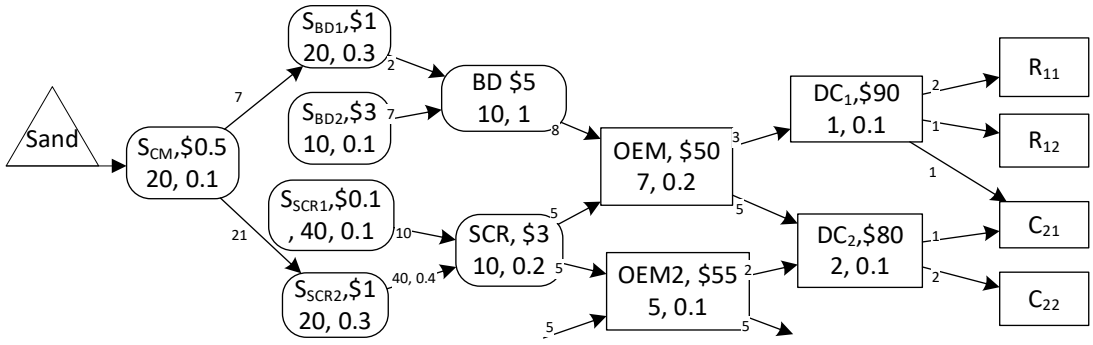


Figure 10.3 Supply network of an electronic product.

5.1 The inventory-positioning model

Let us adapt a tool for analysis in this section. The tool includes a useful symbolic representation of a supply chain with nodes and arcs. A node can be a supplier, an assembly plant, a distribution center, a store, or another facility in a supply chain.

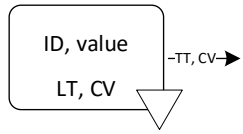


Figure 10.4. Notation for one node and arc in a supply chain.

ID is the node's identifier, such as BD for board supplier. Value is the value of the product at this node. The calculator at OEM would be the calculator's wholesale value. The calculator at a store is LT, its order-fulfillment lead time for parts to go through the node. And CV is the coefficient of variation of the lead time. This can include planning time and production time. If the node is for an end-consumer, LT is the acceptable lead time, $LT_{acceptable}$, and CV is its associated CV. TT is the average transit time in the arc and its CV. If CVs are not listed, the LT or TT are near constant. The triangle indicates inventories. A bucket at the end means that finished goods inventory is carried. A bucket at the beginning of a node means raw material inventory is carried. If there is a number, it means the amount of stock carried in time. These few parameters capture the key information of this node in the supply chain.

5.2 Use inventory positioning model in SC

Figure 10.5 shows a simple supply chain for a tool. It contains a customer C, an assembler A, and a supplier S. The number associated with the arrow is the lead time in transportation. The times are in days.

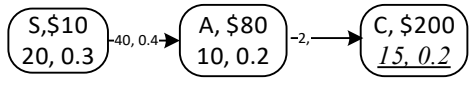


Figure 10.5. A simple supply chain to show the meaning of the labels.

In this example, a pure pull network would mean that when the assembler received the customer order, it will order the component from its supplier in Vietnam. It will take 20 days on average for the supplier to send the order, with $\sigma = 20 * 0.3 = 6$. It will then take 40 days with $CV = 0.4$ to transport to the assembler. The customer's acceptable lead time is 15 days. It is clear that pure pull is not feasible. How can the assembler satisfy the customer's acceptable lead time? The assembler can stock up the completed product at the outbound, or stock up the material from the supplier at the inbound. Which is better?

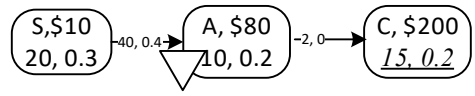


Figure 10.6. Illustration to show the inventory position.

Let us first look at the lead times. If we stock up the assembled product, or at the outbound of the A, we can deliver the product to the customer in 2 days, including both the delivery time and delivery planning time. This is 13 days faster than expected, on average. If we stock up the component from the supplier, or the inbound of A, we can deliver the product to the customer with 12 days on average, with a standard deviation of 10 * 0.2 = 2. If the acceptable leadtime is 15 days, the safety factor for over 15 days is

$$z = \frac{LT_{Acceptable} - \mu}{\sigma} = \frac{15 - 12}{2} = 1.5$$

The probability of less or equal to 15 days is 93.3%. The customer can also tolerate a little variability. Therefore, this is good. However, this is the most upstream point we must position our inventories. Otherwise, we cannot meet the customer's acceptable lead times.

6. HUB-AND-SPOKE NETWORKS

Hub-and-spoke networks make an excellent example to illustrate the economies of competition, scale, sharing, and speed. Many of their features can be extended to analyze the strategies of outsourcing and partnerships.

If you fly from San Francisco to Savannah, you may go from San Francisco to Atlanta, and then on to Savannah. Atlanta and San Francisco are hub cities. Savannah is a spoke. Each day, there are not enough passengers to fly from San Francisco to Savannah. However, there are many people who would fly to Atlanta or other cities in the vicinity. In Atlanta, many passengers from other hub cities or spoke cities may want to go to Savannah. They can be pooled in Atlanta to fill a flight or more to Savannah.

This is the same in the ocean. A 24,000 TEU ship going from Los Angeles to a major hub in Asia can carry many containers to smaller ports near the hub in Asia. Air cargo is the same. Where would you guess is the largest port for air freight in United States? Memphis! That is the air hub for FedEx. Every night, two batches of airplanes arrive, exchange loads, and depart. You can search for a nice video on FedEx high-speed package sorting, and a video tracking of the flights in United States on the internet. Atlanta is one of the major road hubs because there are three interstate highways passing through: I-20, I-75, and I-85. A lot of LTL trucks exchange loads somewhere in Atlanta, including UPS and FedEx, among many others.

6.1 Point-to-point network

Consider that in a region there are *m* suppliers, shown in Figure 10.7 by small squares, and *n* clients, shown by small circles. Each client orders $D(\mu, \sigma^2)$ pallets from each of the suppliers each week. Logically, we can arrange the network in the figure. There are *mn* routes needed to connect. The number of routes in this network is *mn*.

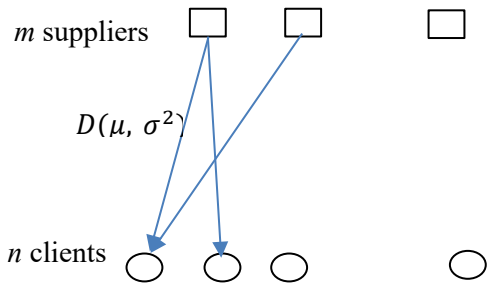


Figure 10.7. A point-to-point delivery network.

6.2 The characteristics of Hub-and-spoke network

Consider the same demand as above but with an added hub, shown in Figure 10.6 as the large rectangle in the center. Let us assume strict hub-and-spoke process: all loads go from a supplier to the hub and then from the hub to the client. The loads are pooled together whenever possible. The flow of goods is as follows in a hub-and-spoke network:

1. The loads from each of *m* suppliers to its all clients are transported to the hub.
2. The loads from all suppliers are sorted for each client.
3. The sorted loads are transported to each of *n* clients.

In a single hub network, all loads are transported to the hub. In the hub, the loads are unloaded and loaded to different trucks destined for their clients.

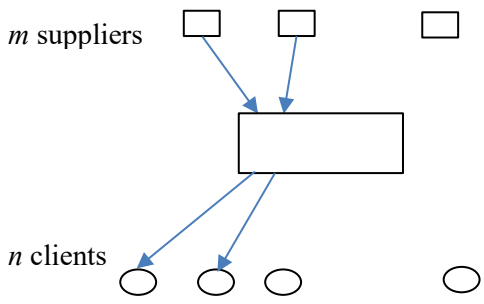


Figure 10.8. A network with a single hub.

6.2.1 Route reduction

The total number of routes in a point-to-point connection was *mn*. With a hub, it is *it is* $m + n$. *If* $m = 100, n = 1000$, then the number of routes reduces from 100,000 to 1,100, a 99% reduction!

6.2.2. The load consolidation from supplier spokes to DC hub.

Since all loads from a supplier must go through the connected hub, the average load from a supplier to the hub is $\mu_S = n\mu$. If the demands are iid and not correlated, the standard deviation and coefficient of variation are

$$\sigma_s = \sqrt{n}\sigma$$

$$CV_s = \frac{\sqrt{n}\sigma}{n\mu} = \frac{1}{\sqrt{n}}\frac{\sigma}{\mu} = \frac{1}{\sqrt{n}}CV$$

The load scaled up by *n* times, and the relative variability decreased by square root of *n*. If *n* = 1000, the scale goes up 1000 times while the relative variability reduces to only about 1/30th. This can go from caseload to truckload and steady.

6.2.3 The sorting at the hub.

At the hub, almost all loads must be unloaded from the trucks. The only possible exception is when a truck has a load to a client and is scheduled to deliver to that client, which was loaded to the truck first, or in the inner most location of the truck. Otherwise, it is in the way of moving the other loads.

If a different set of trucks are used for the delivery to the clients, the loads can be transferred to the delivery trucks directly. This is a pure cross-dock operation. A pure cross-dock facility is a long platform with dock doors on both sides. The forklifts move fast to unload from inbound trucks and to drop off in the outbound trucks. For cases or packages, high-speed conveyors are often used. This will avoid double handling. However, both the inbound trailer and the outbound trailer must be at the dock at the same time, adding a scheduling constraint.

If the inbounds and outbounds are not lined up perfectly, the loads are unloaded, staged, and then loaded to the delivery trucks later. This will require space.

6.2.4 The loads consolidation from DC hub to the client spokes

Similarly, you can find the parameters from the hub to a client. At DC, all loads destined to each client are pooled together. The average load to a client is $\mu_c = m\mu$. Its standard deviation and coefficient variation are

$$\sigma_c = \sqrt{m}\sigma$$

$$CV_c = \frac{\sqrt{m}\sigma}{m\mu} = \frac{1}{\sqrt{m}}\frac{\sigma}{\mu} = \frac{1}{\sqrt{m}}CV$$

If m = 100, the scale goes up 100 times while the relative variability reduces to only one-tenth. This can go from pallet loads to multiple truckloads, and steady.

6.2.5 Pros and cons of hub-and-spoke networks

The hub-and-spoke networks reduce the routes and enjoy the economies of scale and uncertainty pooling. In logistics, the highest cost is in transportation. The unit costs in transportation continue to increase without drastic change, such as self-driving trucks. Therefore, the hub-and-spoke network can be effective in dealing with small and uncertain loads.

However, there is added cost and delay at the hub—two enemies in the supply chain. The cost involves the facilities, equipment, and operations to unload, sort, and load. These take time. However, time is more complicated issue due to the unit load. The loads are also handled at the hub, which can lead to potential damage and errors.

Example

A 3PL services 100 suppliers and 400 clients. Each client demands $D(0.1, 0.3^2)$ mixed pallets per day from each supplier. The 3PL requires the client to put their goods on a pallet for easy loading/unloading. This means there is one pallet load every 10 days. Please find the following.

1. How many pallets should be picked up from each supplier each day?
2. How many pallets should be dropped off at each client each day?
3. What will be the time delays for the supplier to deliver themselves in point-to-point service?
4. Is there opportunity to reduce travel?
5. What will be the time delay in hub-and-spoke service?

Solutions

1. Pick up: $D_s(400 * 0.1, 400 * 0.3^2) = D_s(40, 6^2)$, mixed pallets per day, $CV_s = 0.15$, very low variability each day.
2. Delivery $D_c(10, 3^2)$, mixed pallets per day, $CV_c = 0.33$, low variability. On average this is about one-third of a truckload per day.
3. In a pure point-to-point service, the supplier can send a small truck with one pallet on to one of the 40 clients, on average, that has an order. It will require on average 40 truck-trips per day. It will be the fastest if the supplier has enough trucks and drivers. It will be very costly.
4. The supplier can also provide multi-stop truckload service, or a milk run. It can load the truck in the reverse order of the delivery and drop off the loads as in a milk run. The first stop will be fast, but the last one, around the 40th, can take a long time, depending on many other factors. The supplier can also send more than one truck to speed up the delivery.
5. In the hub-and-spoke operation, the pickup from each supplier is about a truckload with low variability. The truck will only need to visit one supplier each day, and make a single stop at the DC. The delivery truck will require three stops on average. It can be arranged so that the three stops are not too far away. However, all loads are delayed and handled at the hub.

If the 3PL delivers daily, it will require on average three stops on each trip. It can also consolidate to deliver every three days.

6.3 Multi-hub system

In practice, 3PLs may have multiple hubs. Consider a two-hub network, shown in Figure 10.9.

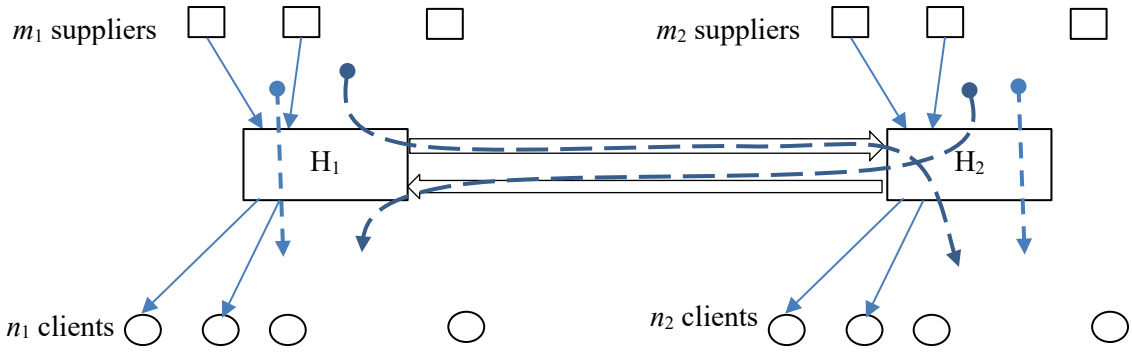


Figure 10.9. A network with two hubs.

First, we assume a strict hub-and-spoke protocol: each spoke is only connected to one hub, and spoke-to-spoke transportation is prohibited. As a result, the inter-hub loads must go through both hubs and through the trunk line connecting the two hubs. Each hub must receive supplies from the directly connected suppliers and the suppliers connected to the other supplier but deliver to the clients connected to itself. There are two flows from hub to hub not directly connected to supplier or client.

1. The total loads from a supplier connected to H_1 to be delivered to clients connected to both H_1 and H_2.

$$\mu_1 = (n_1 + n_2)\mu$$

$$\sigma_1 = \sqrt{n_1 + n_2}\sigma$$

$$CV_1 = \frac{1}{\sqrt{n_1 + n_2}} CV$$

2. The total load from to a client connected to H_1.

$$\mu_1 = (m_1 + m_2)\mu$$

$$\sigma_1 = \sqrt{m_1 + m_2}\sigma$$

$$CV_1 = \frac{1}{\sqrt{m + m_2}} CV$$

3. The expected total loads from H_1 to H_2.

$$\mu_{H1-H2} = n_2 m_1 \mu$$

4. Total average loads H_1 must sort.

$$\mu_{H1} = [m_1(n_1 + n_2) + m_2 n_1]\mu$$

5. The load from all suppliers connect to hub 1 to a single client connected to hub 2.

$$\mu_{H1} = m_1 \mu$$

This quantity is useful when this load is nearly sufficient to be a truckload. In that case, the truck can bypass hub 2 and directly go to the client. UPS and FedEx do this often. Within a few hundred miles, the truck delivery be as fast as air. You can imagine the amount of savings.

6.4 Revisit to in-house and outsource strategies

Now that we have discussed all the competitive and some collaborative strategies, we can revisit in-house versus outsourcing.

The main benefits of adding or keeping in-house supply chain functions help to grow and to develop competence. Most of all, it helps with direct control. Others may not put your service as high priority. Even if they do, they may not have place the same priority as you would place among your clients.

The main benefits of outsourcing, in addition to service provider competency in the domain it serves, are in the economies of scale, sharing, uncertainty pooling, and possibly greater speed in normal situations.

7. NETWORK DESIGN AND HUMAN NEEDS

The objective of network design is to minimize cost while maintain acceptable speed. The major forces in cost reduction are economies of scale and uncertainty reduction. Their linkage to human needs has already been discussed in the earlier chapters.

11. QUANTITATIVE INCENTIVE SUPPLY CHAIN CONTRACTS

People and organizations strive towards incentives. In this section, we will explore the quantitative incentives related to inventories. In Chapter 3, we presented the principal-agent models and a 2-level special case with high pay and low pay to incentivize the agent to commit high effort before and after the contracts. In practice, the incentives can be in many dimensions, such as incentives in transportation changes or non-financial incentives such as awards in the waste diversion project covered in Chapter 2. They can also be related to quantity, contract terms, quality, speed, preferential treatment, and other things.

Supply networks are complicated. Often, they can be partitioned into a two-tier supply chain building block. The pair can be a manufacturer and an assembler, an assembler and a dealer, a distributor and a retailer, a lab and a hospital, or a professor and her students.

1. TWO-TIER SUPPLY CHAIN MODEL

Two supply partners can be considered a supply-chain pair: a supplier and a client, shown in Figure 11.1. We will use this model to explore the use of incentives to increase profit for both partners.

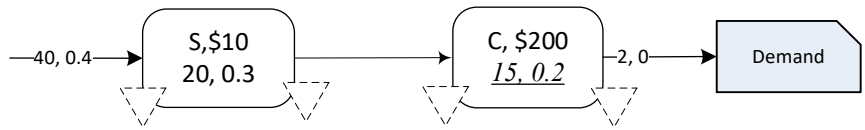


Figure 11.1 The possible inventory locations between a supplier and a client.

1.1 Best location to keep inventory

The demand is normally uncertain. The customer's acceptable lead time is normally shorter than the lead time from the beginning of the supplier to the customer. Therefore, inventory must be kept somewhere in the system to remove the lead time before the inventory point. Since the inventories are built before the demand is known, someone must take the risk of the inventory, either too much or too little.

In this simple pair, there are four possible inventory locations: before or after the supplier, or before or after the client. Two of the most important decisions for the inventory are the location and ownership. The inventory cost increases as one moves downstream in the supply chain, because of value added operations and because of less flexibility in the product. In a specific market, the acceptable lead times are established in the market. Therefore, the inventory should be at the most upstream position and satisfy the acceptable lead times.

Consider the fast-fashion market between a producer and a retailer. For fashion, the setup cost in design, marketing, and transportation is high. The batch size is normally large. The demand for high fashion is highly uncertain. The customer's acceptable lead time from the client is short. When someone wants a new pair of jeans, he or she wants it quickly. The customer also wants to try the jeans on before purchasing. However, the design, production, and transportation require long lead times. There must be inventory stored at the retailer, as shown in Figure 11.2.

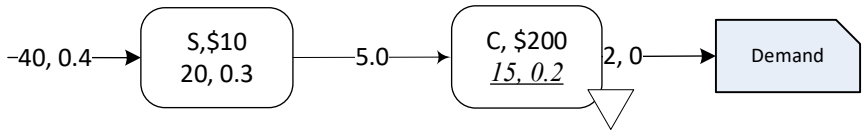


Figure 11.2. When the finished good inventory must be stored at the client's location.

1.2 Risk owner

The ownership of the risk associated with an inventory depends on the market dynamics and a player's power determined in Michael Porter's five forces of competition—market position, availability of alternatives, financial security, barrier to entry, and saturation.

In the high-fashion example, if the supplier is a well-established strong brand, and has many retailers to choose from, a weak retailer has to take the risk to order before the demand is known.

On the other hand, if the product is a commodity type with many potential suppliers, and the retailer is a well-established chain with a large customer base, the producer may have to take the risk to make-to-stock and let the retailer order from already committed inventories, shown in Figure 11.3.

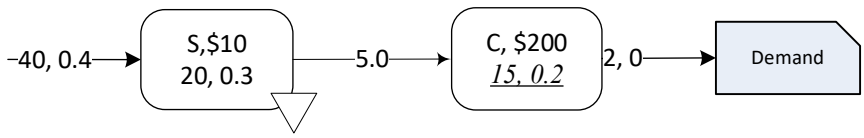


Figure 11.3. When the supplier must keep finished goods as inventory.

An extreme example where a client dominates is Dell Computer in the 1990s. Dell was an online custom-made computer supplier. Dell required its suppliers to keep sufficient inventory on Dell's site. When an online order was received, Dell would plan its production. Once the production of a customer order started, Dell would scan the CPU, memory, HD, and so on needed for a particular order to form an assembly kit. The scan changed ownership of the parts from the supplier to Dell. It would take an hour or two for Dell to complete the assembly after kitting, and about a day for testing. The time from the customer payment to the time the custom computer

was shipped was only two to three days. Dell had almost 30 days to pay the suppliers on the parts, so Dell owned zero risk while the supplier owned 100% of the inventory risk. What a great deal for Dell! Some people say that Dell carried a negative inventory because it got payment from online buyers almost four weeks before it paid its suppliers. The inventory position can be represented in Figure 11.4.

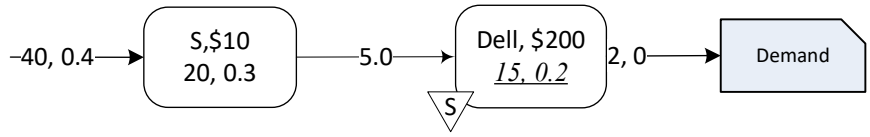


Figure 11.4. Dell Computer's inventory arrangement with its suppliers.

There can be many other possible arrangements between partners in practice. Amazon Marketplace is an arrangement between Amazon.com and retailers that sell products using Amazon's platform. How do you model the relationship between the supplier and Amazon?

1.3 When the dominant partner does not own the inventory risk

There are four major possible dominance and inventory risk combinations:

1. The supplier dominates but the client owns the inventory risk.
2. The supplier dominates and also owns the inventory risk.
3. The client dominates but the supplier owns the inventory risk (such as with Dell).
4. The client dominates and also owns the inventory risk.

In combinations 1 and 3, the incentives are misaligned. The dominant player could enjoy risk-free business, such as what Dell did initially. We will see in the following sections that in such situations the dominant player can help to grow the pie for both partners. This is the way Dell changed its practice later.

In combinations 2 and 4, the incentive is aligned.

2. THE SUPPLIER DOMINATES BUT THE CLIENT OWNS INVENTORY RISK

This often happens in the make-to-order market. The client preorders the goods from the supplier in a predetermined quantity, Q, and takes the risk of excess inventory or shortage. In high fashion or other custom-made products, the make-to-order market is common. The make-to-order model can also be true when the upstream player is strong and there are many downstream players competing for the business. In such situation, the downstream player has to take the risk to determine the order quantity before the uncertain demand, because the transaction price/cost between the two players can share the same notation and cause confusion.

2.1 No incentive contract

We first develop the profit model without incentive contract. Let

- p_o be the price of a product selling out to its customer or client of a player.

- p_i be the price of a product buying in by a player.
- p_s be the price of salvage value.
- Q be production or order quantity for the risk-taker, a decision variable.

where $p_o > p_i > p_s$

Example

A well-known brand tool-maker produces a special seasonal tool useful in a region. It sells these tools through dealers. In this relationship, the manufacturer, or the tool-maker, can select from many competing dealers. The wholesale price is \$700, and the retail is \$1000. The salvage is \$100. The production cost is \$400, with setup cost of \$5,000. The demand in a season for one of the dealers can be approximated by $D \sim N(300, 100^2)$. The questions are:

1. Should the dealer determine order quantity or the tool-maker determine production quantity?
2. How much should the quantity be?
3. How much is dealer profit?
4. How much is the manufacturer's profit?

Since the tool-maker has options, it is up to the dealers to take the initiative to bid for the opportunity to sell the tools. Therefore, the dealer has to take the risk to order before the season. The dealer has to determine the order quantity, Q, and order from the tool-maker. The tool-maker will manufacturer the amount ordered without risk. The answers to the next three questions are more involved. We will derive these in the next section.

2.2 Newsvendor model for profit

In Chapter 9, we used the newsvendor model to model the cost associated with uncertainty. In this chapter, we need to develop the newsvendor model for profit.

If the actual demand realized at the dealer is $x \leq Q$, shown in Figure 11.5, it will sell quantity x at retail price. In addition, there will be quantity $Q - x$ unsold units left to be salvaged at a salvage price. If the demand quantity $x > Q$, it can only sell Q, or what it has ordered, with a loss of potential profit of amount it did not order. There is no salvage in this case. This is a newsvendor problem with the objective function of maximizing profit.

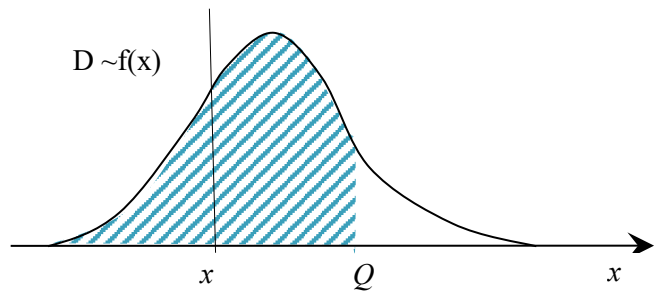


Figure 11.5. Illustration of the order quantity and demand distribution.

2.1.1 Optimum order quantity without contract

The optimum order quantity is determined by the critical ratio. In Chapter 9, we derived the formula for minimum order quantity and its cost. In this section, we need to compare the profit with different contracts. Therefore, we must develop the formula for maximum profit order quantity and the profit.

$$\begin{aligned}\Pi(Q) &= p_o E(\text{Units Sold}) + p_s E(\text{Units unsold}) - p_i E(\text{Units paid for}) \\ &= p_o\left[\int_{-\infty}^{Q} xf(x)dx + \int_{Q}^{\infty} Qf(x)dx\right] + p_s\int_{-\infty}^{Q}(Q-x)f(x)dx - p_i\,Q\end{aligned}$$

We can apply the first order condition to find the optimum order quantity. Since the decision variable is also in the limits, you can apply the Leibniz rule.

$$\begin{aligned}\frac{d\Pi(Q)}{dQ} &= p_o\frac{d}{dQ}\left(\int_{0}^{Q} xf(x)dx + \int_{Q}^{\infty} Qf(x)dx\right) + p_s\frac{d}{dQ}\int_{0}^{Q}(Q-x)f(x)dx - p_i \\ &= p_o\left[0 + Qf(Q) - 0 + \int_{Q}^{\infty} f(x)dx + 0 - Qf(Q)\right] + p_s\left(\int_{0}^{Q} f(x)dx + 0 - 0\right) - p_i \\ &= p_o\left[1 - F(Q)\right] + p_o F(Q) - p_i \\ &= p_o - p_i - (p_o - p_s)F(Q) \to 0\end{aligned}$$

Since $p_o > p_s$, the second derivative is less than 0, the solution Q yields maximum profit. Rearrange the terms, we have

$$F(Q^*) = \frac{p_o - p_s}{p_o + p_s}$$

You may notice that this is the same form as the critical ratio in section 3. Is this a coincidence? Actually, you can consider underage cost as the lost profit, or $c_u = p_0 - p_i$, and the overage cost is the cost of paying for the product minus salvage, or $c_o = p_0 - p_s$. They are the same! Note that salvage price can be negative. Since $p_o > p_i > p_s$, this ratio is between $0 < F(Q^*) < 1$.

Example 2

In this example, we have

$$F(Q^*) = \frac{p_o - p_i}{p_o - p_s} = \frac{1000 - 700}{1000 - 100} = 0.3333$$

This is a short but implicit solution. It applies to any continuous distribution. For the discrete distribution, you can prove the equal sign is changed to ≤. If the demand distribution can be approximated by normal distribution, we need to find Q by using the table for standard normal distribution. You can find the corresponding z value as $z = \Phi^{-1}\big(F(Q^*)\big)$. You can also find it using the Excel function for standard normal NORMSINV(critical ratio), or for a generic normal distribution NORMINV(critical ratio, 0, 1), or on your calculator. The final solution for the generic normal will be

$$Q = \mu + \sigma z$$

where z is called safety factor because it is the proportional parameter for the safety stock. If there is also a cost of good will, where will that go? It is associated with the number of units short, or

$$C(Q) = c_g \int_Q^\infty (x - Q) f(x) dx$$

We will revisit this later.

2.3 Optimum profit for each partner without contract

If the demand is normal, we can find the profit as follows.

$$\begin{aligned}\Pi(Q) &= p_o \int_{-\infty}^{Q} x f(x) dx + p_o \int_{Q}^{\infty} Q f(x) dx + p_s \int_{-\infty}^{Q} (Q - x) f(x) dx - p_i Q \\ &= p_o \left[\int_{-\infty}^{Q} x f(x) dx + Q\big(1 - F(Q)\big) \right] + p_s \left[QF(Q) - \int_{-\infty}^{Q} x f(x) dx \right] - p_i Q\end{aligned}$$

If we can solve the remaining integral, we can find the closed-form solution. For normal distribution, $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} = \frac{1}{\sigma}\varphi\left(\frac{x-\mu}{\sigma}\right)$. We can find everything directly or via a normal table, except $\int_{-\infty}^{Q} x f(x) dx$. Even for this one, we can find the solution as follows.

$$\begin{aligned}
\int_{-\infty}^{Q} xf(x)dx &= \int_{-\infty}^{Q} x\frac{1}{\sigma}\phi\left(\frac{x-\mu}{\sigma}\right)dx \\
&= \int_{-\infty}^{\frac{Q-\mu}{\sigma}} (\mu+\sigma z)\frac{1}{\sigma}\phi(z)\sigma dz \\
&= \int_{-\infty}^{\frac{Q-\mu}{\sigma}} (\mu+\sigma z)\phi(z)dz \\
&= \mu\Phi\left(\frac{Q-\mu}{\sigma}\right)+\sigma\int_{-\infty}^{\frac{Q-\mu}{\sigma}} z\phi(z)dz \\
&= \mu F(Q)-\sigma\phi\left(\frac{Q-\mu}{\sigma}\right) \\
&= \mu F(Q)-\sigma\phi(z)
\end{aligned}$$

Then, the profit is

$$\begin{aligned}
\Pi_{Down}(Q) &= p_o\left[\int_{-\infty}^{Q} xf(x)dx + Q(1-F(Q))\right] + p_s\left[QF(Q)-\int_{-\infty}^{Q} xf(x)dx\right] - p_iQ \\
&= p_o\left[\mu F(Q)-\sigma\varphi(z)+Q-QF(Q)\right]+p_s\left[QF(Q)-\mu F(Q)+\sigma\varphi(z)\right]-p_iQ \\
&= p_o\left[Q+(\mu-Q)F(Q)-\sigma\varphi(z)\right]+p_s\left[(Q-\mu)F(Q)+\sigma\varphi(z)\right]-p_iQ \\
&= p_oE\left[\text{Number sold}\right]+p_sE\left[\text{Number not sold}\right]-p_i\left[\text{Number ordered}\right]
\end{aligned}$$

The result is long, but they are either known or available on the normal table.

Examples 3

Now we can answer the third and fourth parts of the example.

For the example problem from part 2, we have

$$F(Q^*)=\frac{p_o-p_i}{p_o-p_s}=\frac{1000-700}{1000-100}=0.3333$$

Note that the calculations are based on Excel, not the tables with truncation. They are slightly different.

$z = \Phi^{-1}(0.3333) = -\Phi^{-1}(1 - 0.3333) = -0.4307.\ \varphi(-0.4307) = 0.364.$

$Q = \mu + \sigma z = 300 - 100 * 0.4307 = 256.9 \sim 257.$

This should be rounded up to 257.

$E[\text{Sold}] = Q + (\mu - Q)F(Q) - \sigma\varphi(z)_S = 257 + 43 * 0.333 - 100 * 0.364 = 234.9$

$E[\text{Unsold}] = (Q - \mu)F(Q) + \sigma\varphi(z) = -43 * 0.333 + 100 * 0.364 = 22.03$

$\Pi_{Down}(257) = 1000 * 234.9 + 100 * 22.03 - 700 * 257 = \$57{,}276$

$\Pi_{Up}(257) = 700 * 257 - 400 * 257 - 5000 = \$72{,}100$

The total profit for both players combined is $129,376.

2.4 Revenue-sharing contract

Note that without collaboration, the expected demand is 300. The order size is only 257, 43 units below the expected value. The reason is that the profit lost from the shortage is $300 per tool. However, the cost of excess is 600, much higher than the profit gain. Therefore, the dealer tends to order less than the expected demand.

One way to collaborate is to share some the risk with dealers by offering some incentive for the dealer to order more, such as a discount on the ordering cost $p_{discount}$, and in return receive part of the revenue after sales p_{share}. In this arrangement, the dealer will take less risk because the extra cost, or the share returned to the tool-maker will take place after the sales. If the discount is equal to the revenue share, there can be benefits to both. Even if the share is higher than the discount, the dealer may still like the contract, because the share is from its profit already secured. Let us consider the equal amount of $100 first. To get a better understanding of the exchanges, let us start from the original model

$$\begin{aligned}\Pi_{Down}(Q) &= p_o E[\text{Sold}] - p_{Share} E[\text{Sold}] + p_s E[\text{Unsold}] - (p_i - p_{Discount})Q \\ &= (p_o - p_{Share})E[\text{Unit Sold}] + p_s E[\text{Unit not sold}] - (p_i - p_{Discount})Q\end{aligned}$$

We can substitute the quantities in the critical ratio, term by term, and have

$$F(Q^*) = \frac{p_o - p_i}{p_o - p_s} = \frac{(p_o - p_{share}) - (p_i - p_{discount})}{(p_o - p_{share}) - p_s} = \frac{(1000 - 100) - (700 - 100)}{(1000 - 100) - 100} = \frac{300}{800} = 0.375$$

$\Phi^{-1}(0.375) = -\Phi^{-1}(1 - 0.375) = -0.3186,\ \varphi(-0.3186) = 0.379$

$Q = \mu + \sigma z = 300 - 100 * 0.3186 = 268.1 \rightarrow 269.$

$E[\text{Sold}] = Q + (\mu - Q)F(Q) - \sigma\varphi(z) = 269 + 31 * 0.375 - 100 * (0.379) = 242.7$
$E[\text{Unsold}] = (Q - \mu)F(Q) + \sigma\varphi(z) = -31 * 0.375 + 100 * 0.379 = 26.3$

$\Pi_{Down}(269) = (1000 - 100) * 242.7 + 100 * 26.28 - 600 * 269 = \$59{,}660$ (vs. \$57,276, 4% reduction)

$\Pi_{Up}(269) = (600 - 400) * 269 - 5000 + 100 * 242.7 = \$73{,}070$ (vs. \$72,100, 1.3% reduction)

Both benefit from the revenue share as a form of risk-sharing, although the retailer gained more. The reason is that the total expected sales are higher.

The manufacturer can ask for a higher share, say \$105. Without more advanced techniques, you can use numerical methods through algorithm and search.

2.5 Buy-back contract

The reason that the downstream player ordered 14% below the expected demand is that the overage cost of 600 is much higher than the loss of profit, or underage cost of 300. The smaller order quantity can potentially hurt the manufacturer as well because his total profit depends on the quantity ordered.

If the manufacturer offers to buy back unsold units at some price higher than salvage, the dealer may order more, and therefore there will be higher total revenue and possibly higher profit for both. The manufacturer should provide between salvage and wholesale because less than salvage value, the dealer would just salvage the leftover tools, and at higher than wholesale, the dealer would have no incentive to sell.

The tools can be worth more than salvage value to the tool-maker for the following reasons. The tool-maker can:

1. Buy back the tools and sell to different regions. The production costs the manufacturer \$400 for each tool, plus the share of the setup cost. If there are shortages in other markets, the tools bought back can be used in other markets to avoid additional production cost. The value will be the production cost plus the share of the setup cost minus handling cost (shipping, handling, bookkeeping).
2. Store them to satisfy future demands. The dealers are often at a good location with well-appointed showrooms at high cost. The tool-maker does not have to be at a good location and does not have to face customers. If the tool does not change quickly over time, they will still be as valuable later. The cost of storage is lower than that for the dealer. In

addition, if the tools going to the salvage, they will most likely enter the secondary market. The sales in the secondary market can erode or cannibalize the demand for the new tools in the future. The secondary market can also tarnish the brand.

Let us consider the first case. The manufacturer can value the unsold tool at production cost of $400 per tool because it does not have to make it again. There is additional cost of $100 for the manufacturer to account for picking up, inventory, restocking, and so on.

$$F(Q*) \geq \frac{p_o - p_i}{p_o - p_s} = \frac{1000-700}{1000-400} = 0.5$$

Therefore, order the average amount. The profit for the distributor is

$$z = 0,\ \varphi(0) = 0.399$$

$$Q* = 300 + 100 * 0 = 300$$

$$E[\text{Sold}] = Q + (\mu - Q)F(Q) - \sigma\varphi(z)_s = 300 + 0 - 100 * .399 = 260.1$$

$$E[\text{unsold}] = (Q - \mu)F(Q) + \sigma\varphi(z) = 39.9$$

$\Pi_{Down}(300) = 1000 * 260.1 + 400 * 39.9 - 700 * 300 = \$66.060$ (vs. $57,276 before)

$\Pi_{Up}(300) = (700 - 400) * 300 - 5000 + (-400 + 400 - 100) * 39.9 = \$81,010$ (vs. $72,100)

Both players gain from the buy-back contract. The extra profit is generated from the extra expected sales, and the reclaimed value from the unsold units. The first −400 in the parentheses is the pay to the dealer, the +400 is the value of tools bought back, and −100 is the restocking fee.

If the manufacturer makes the tool-dealer absorb the restocking fee of $100, the manufacturer will only pay $300 to buy the tools back. $300 is now the new salvage value for the dealer. His critical ratio is now

$$F(Q*) = \frac{p_o - p_i}{p_o - p_s} = \frac{1000-700}{1000-300} = 0.4286$$

Therefore, order the average amount. The profit for the downstream distributor is

$$z = \Phi^{-1}(0.4286) = -\Phi^{-1}(1 - 0.4286) = -0.1800,\ \varphi(-0.4286) = 0.393$$

$Q* = 300 + 100 * (-0.180) = 282$

$E[\text{Sold}] = Q + (\mu - Q)F(Q) - \sigma\varphi(z)_s = 300 + 18 * 0.4286 - 100 * 0.394 = 250.5$

$E[\text{unsold}] = (Q - \mu)F(Q) + \sigma\varphi(z) = 18 + 100 * 0.393 = 31.5$

$\Pi_{Down}(282) = 1000 * 250.5 + 300 * 31.5 - 700 * 282 = \$62{,}550$ (vs. \$57,276 before)

$\Pi_{Up}(282) = (700 - 400) * 282 - 5000 + (-300 + 400) * 31.54 = \$82{,}754$ (vs. \$72,100)

Both players gain from the buy-back contract. The extra profit is generated from the extra expected sales, and the reclaimed value from the unsold units. The first −400 in the parentheses is the pay to the dealer, the +400 is the value of tools bought back, and 300 is paid to the dealer by the manufacturer (400 − 100 restocking fee).

3. PERFECT COLLABORATION OR GLOBAL OPTIMUM SOLUTION

If the manufacturer owns the dealer, or the two form a perfect alliance, meaning they would only consider the total profit for both, not the individual profit, then the two would make decisions based on the external demand but not the transactions between them. This is basically a vertical integration, or merging of the two firms. The vendor-managed inventory (VMI) can also achieve the similar results.

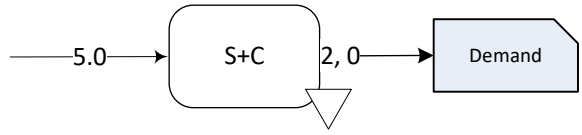


Figure 11.6. When the supplier and client have perfect collaboration with inventories.

The ordering from the dealer to the manufacturer is immaterial. In this case, the cost is the production cost. Therefore, we have

$F(Q*) = \frac{p_o - p_i}{p_o - p_s} = \frac{1000 - 400}{1000 - 100} = 0.667$

$z = 0.4307,\ \varphi(z) = 0.374$

$Q^* = 300 + 100 * 0.4307 = 343.07.$

Theoretically, you should round this to 344. However, the profit difference between 344 and 343 is very small, if you do everything correctly because the z value, density, and distribution at this value is different from the critical ratio. Let us try 344.

$$E[\text{Sold}] = Q + (\mu - Q)F(Q) - \sigma\varphi(z) = 344 + (300 - 344) * .667 - 100 * .374 = 277.25$$

$$E[\text{unsold}] = (Q - \mu)F(Q) + \sigma\varphi(z) = 44 * .667 + 100 * 0.374 = 66.75$$

$$\Pi_{Combined}(344) = 1000 * 277.25 + 100 * 66.75 - 400 * 344 - 5000 = \$141{,}325$$

$$\Pi_{Combined}(343) = 1000 * 276.92 + 100 * 66.08 - 400 * 343 - 5000 = \$141{,}328$$

(Even Q = 343 is slightly higher, but this is due to rounding errors).

This total profit among the two is 10% higher than no contract, 7.4% higher than revenue share contract, but 6% lower than the buy-back contract. The reason this is lower than the buy-back contract is that in the buy-back contract, we value the salvage much higher. If we do the same here, the global optimum solution will give you even higher profit.

4. THE CLIENT DOMINATES BUT THE CLIENT OWNS INVENTORY RISK

If the downstream dominates, the upstream player has to take the risk to make the product and store them in stock, or procure to stock. This is similar to make-to-stock market. Many commodity products are in this type of market. The profit margin is often lower with such products. Let us consider an example in which the manufacturer is supplier to large and powerful dealers. The dealers demand short lead time and lower prices. The manufacturer must make-to-stock. Let us consider the same set of prices and costs.

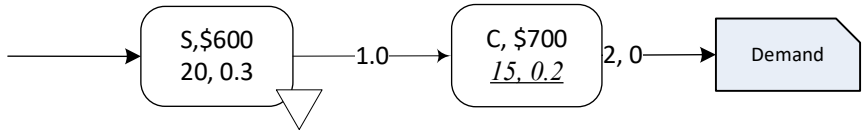


Figure 11.7. Make-to-stock inventory situation.

4.1 No contract

Let us consider a commodity tool that the manufacturer produces for the large and powerful dealers who dominate the manufacture. The retail price and wholesale price are the same, $1000 and $700, respectively. However, it costs the manufacturer $600 to make and deliver the tool. We assume that the demand follows the same distribution for one dealer.

$F(Q^*) = \frac{p_o - p_i}{p_o - p_s} = \frac{p_o - c}{p_o - p_s} = \frac{700-600}{700-100} = 0.1667$

$\Phi^{-1}(0.1667) = -\Phi^{-1}(1 - 0.1667) = -0.967.$

$Q = \mu + \sigma z = 300 - 100 * 0.967 = 203.3 \sim 204$, 68% of the average demand.

$E[\text{Ordered}] = Q + (\mu - Q)F(Q) - \sigma\varphi(z)_s = 204 + 96 * 0.1667 - 100 * 0.250 = 195.00$

$E[\text{Unordered}] = (Q - \mu)F(Q) + \sigma\varphi(z) = -96 * 0.1667 + 100 * 0.250 = 8.997 = 9.00$

$\Pi_{Supplier}(204) = 700 * 195 - 600 * 204 + 100 * 9 - 5000 = \$10{,}000$

$\Pi_{Dealer}(204) = (1000 - 700) * 195 = \$58{,}500$

The dealer makes a lot, but the manufacturer only makes very little. This is as expected, because the dealers dominate the manufacturer.

4.2 Pooling effect impact in one-to-many supply tree

In reality, a manufacturer produces commodities for many dealers. It can profit by the economies of scale. The multiple dealers' random demand may be independent or even negatively correlated. As a result, the manufacturer will also enjoy the pooling effect. Let us assume that the manufacturer is able to expand the market to nine dealers. How do you quantify the profit because of the expansion?

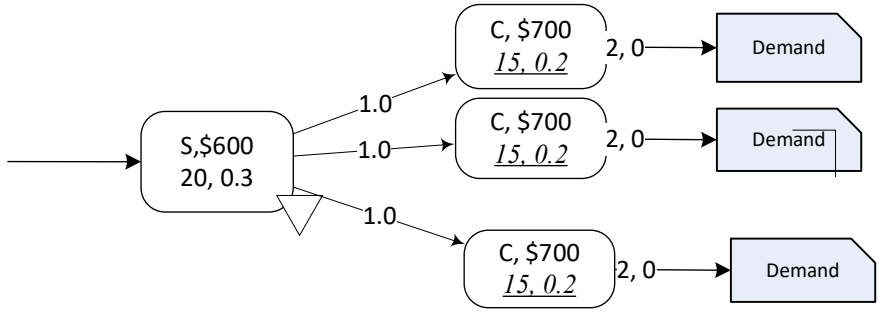


Figure 11.8. The demand pooling effect.

If we assume that the demands at all nine dealers are independent and identically distributed (iid) with the same mean and variance, and the demand at the nine dealerships not correlated. Then the total demand to the supplier is the sum of nine distributions, or $N(2{,}700, 300^2)$.

If the production, salvage, and wholesale prices stay the same, the critical ratio is the same. The total production will be

$Q = \mu + \sigma z = 2700 - 300 * 0.967 = 2409.1 \sim 2410$, 80% of the average! 12% higher.

$E[\text{Ordered}] = Q + (\mu - Q)F(Q) - \sigma\varphi(z)_s = 2410 + 290 * 0.1667 - 300 * 0.250 = 2383.34$

$E[\text{Unordered}] = (Q - \mu)F(Q) + \sigma\varphi(z) = -290 * 0.1667 + 300 * 0.250 = 26.66$

$\Pi_{Supplier}(203) = 700 * 2383.34 - 600 * 2410 + 100 * 26.66 - 5000 = 220{,}004$ (vs. 100,000, more than doubled!)

How to find a dealer's profit? The dealers will share the same inventory pool, either until they are sold out or cannot sell. If we assume dealers do not have any inventory left at the end of the season, we can find the combined profits for all nine dealers.

$\Pi_{All\ dealers}(2410) = (1000 - 700) * 2410 = \$723{,}000$, vs. $58{,}500 \times 9 = 526{,}500$, 37% increase.

Not all dealers will have the same profit.

4.3 Payback contract

Note that without collaboration, the expected number sold is less than 70% of the expected demand. That means that the dealer can do much better if the manufacturer produces more. To achieve it, the dealer can offer some incentive for the manufacturer to produce more. Similar to the buy-back contract, the dealer can offer a payback contract. In a payback contract, the dealer agrees that for each tool *produced but not ordered*, it will offer some money back.

Let us first try a \$100 payback. It means that the dealer, or the downstream player, would pay the manufacturer, or upstream player, \$100 for each tool that was made but not ordered. For the manufacturer, the \$100 payback functions as the salvage value, because it applies to each unit produced but not ordered, therefore,

$F(Q^*) = \frac{p_o - c}{p_o - (p_{Payback} + p_s)} = \frac{700 - 600}{700 - (100 + 100)} = 0.2$

$z = \Phi^{-1}(0.2) = -\Phi^{-1}(1-0.2) = -0.842.$

$Q = \mu + \sigma z = 300 - 100 * 0.842 = 215.84 \sim 216$

$E[\text{Sold}] = Q + (\mu - Q)F(Q) - \sigma\varphi(z)_s = 216 + 84 * 0.2 - 100 * 0.28 = 204.80$

$E[\text{Unordered}] = (Q-\mu)F(Q) + \sigma\varphi(z) = -84 * 0.2 + 100 * 0.28 = 11.196$

$\Pi_{Up}(216) = 700 * 204.8 - 600 * 216 + 200 * 11.196 - 5000 = \$11{,}196$ (vs. \$10,000)
$\Pi_{Down}(216) = (1000 - 700) * 204.80 - 100 * 11.196 = \$60{,}322$ (vs. \$58,254)

By offering incentive for the manufacturer to produce more, the dealer profits more, and so does the manufacturer. Is \$100 payback the best? We picked the number from thin air. Since the order quantity is still significantly lower than the expected demand, you may think that a higher payback amount might be more beneficial. It is true. If the dealer offers a \$200 payback, the profits for both will be \$12,289 and \$62,386. Should the dealer go higher or lower than \$200? Does an optimum value exist? How do you find the optimum value?

As we have seen earlier, the supplier will actually enjoy the pooling effect among multiple dealers and make more. The dealers will also make more.

4.4 Cost-share contract

As opposed to a revenue-share contract in the seller's market, a cost-share scheme can be used in a buyer's market. Here, the upstream player has to take the risk without contract. The downstream player can incentivize the suppler to increase supply quantity by *sharing the production cost of the unit not ordered. In return, the dealer will receive a discounted ordering cost*. For example, if the dealer offers to share \$100 for each tool produced, and in return to receive \$100 discount on the wholesale price, the critical ratio will be

$$\begin{aligned}\Pi_{Up}(Q) &= p_o E[\text{ordered}] - p_{Discount} E[\text{Ordered}] + p_s E[\text{Unordered}] - (c - p_{Share})Q \\ &= (p_o - p_{Discount})E[\text{Unit Sold}] + p_s E[\text{Unit not sold}] - (c - p_{Share})Q\end{aligned}$$

$F(Q^*) = \frac{p_o - p_i}{p_o - p_s} = \frac{(p_o - p_{Discount}) - (c - p_{Share})}{(p_o - p_{Discount}) - p_s} = \frac{(700-100)-(600-100)}{(700-100)-100} = 0.2$

The production quantity, the ordered and unordered quantities are the same as before.

$z = \Phi^{-1}(0.2) = -\Phi^{-1}(1-0.2) = -0.842.$

$$Q = \mu + \sigma z = 300 - 100 * 0.842 = 215.84 \sim 216$$

$$E[\text{Ordered}] = Q + (\mu - Q)F(Q) - \sigma\varphi(z)_s = 216 + 84 * 0.2 - 100 * 0.842 = 204.80$$

$$E[\text{Unordered}] = (Q - \mu)F(Q) + \sigma\varphi(z) = -84 * 0.2 + 100 * 0.842 = 11.196$$

$$\Pi_{Up}(216) = (700 - 100) * 204.80 - (600 - 100) * 216 + (100 + 100) * 11.196 - 5000$$
$$= \$12{,}119 \text{ (vs. } \$10{,}009)$$

$$\Pi_{Down}(216) = (1000 - 700) * 204.80 - 100 * 11.196 = \$60{,}322 \text{ (vs. } \$58{,}254)$$

Although the mechanisms are different, the values are the same as the payback contract. The money exchange also takes place at different times.

5. DISCRETE DEMAND DISTRIBUTION

In many cases, the demand distribution can be discrete. In discrete part orders, we have already practiced rounding up the order quantity. However, the discrete interval can be much wider, such as 100s, 1000s, and so on. This can be due to the unit quantity or lumping of the demand record. For example, the unit load can be in cases of 100 each, or in pallets of 3,000 each. In such cases, the result must be the multiples of the unit quantity. Another example can be the customer demands are in multiples of 100, the critical ratio gives a result between 300 and 400, so what should you order? It is also possible that the historical record of demands is in the 1000 buckets. It can be proved that the result should be

$$P(Q^*) \geq \frac{p_o - p_i}{p_o - p_s}$$

Intuitively, any amount between 300 and 400 cannot be good because the extra will have nowhere to go. The question is why would $\Pi(400) \geq \Pi(300)$? They can be very similar but the larger ordering quantity always dominates the lower ordering quantity.

In discrete distribution, we substitute the integration with summation, and PDF with mass function.

$$G(Q) = p_o E(\text{Units Sold}) + p_s E(\text{Units not sold}) - p_i E(\text{Units paid for})$$
$$= p_o \left[\int_{-\infty}^{Q} x f(x) dx + \int_{Q}^{\infty} Q f(x) dx \right] + p_s \int_{-\infty}^{Q} (Q - x) f(x) dx - p_i Q$$
$$= p_o \left[\sum_{D_i < Q} D_i p(D_i) + Q \sum_{D_i \geq Q} p(D_i) \right] + p_s \left[\sum_{\forall D_i < Q}^{Q} (Q - D_i) p(D_i) \right] - p_i Q$$

The solution for order quantity is the same critical ratio, but substitute "=" with "≥".

Example

The demand is given as in the table.

D	*p*	*P*
8,000	0.110	0.110
10,000	0.110	0.220
12,000	0.285	0.505
14,000	0.220	0.725
16,000	0.175	0.900
18,000	0.100	1.000

The retail price is $125. The price that the retailer pays the manufacturer is $80. The retailer's salvage value is $20. The cost of manufacturing is $35, with a fixed cost of $100,000. Please find the following.

1. Retailer's best ordering quantity
2. Retailer's profit
3. Manufacturer's profit

Solutions: Make-to-order

1. Use the critical ratio to find the value of the distribution.

$$F(Q^*) \geq \frac{p_{retail} - p_{wholesale}}{p_{retail} - p_{salvage}} = \frac{125-80}{125-20} = 0.4286$$

Based on the cumulative probability, the first demand level with a probability greater than 0.429 is 12,000.

The expected number of units sold if Q = 12000. It is best to do the sum separately.

$$E(\text{Sold}) = \sum_{\forall D_i < Q}^{Q} D_i p(D_i) + Q \sum_{\forall D_i \geq Q}^{\infty} p(D_i)$$
$$= (8000 * 0.11 + 10000 * 0.11) + 12000 * (0.285 + 0.220 + 0.175 + 0.100)$$
$$= 1980 + 9360 = 11340$$

The expected number of units not sold or going to salvage.

$$E(\text{Unsold}) = \sum_{\forall D_i < Q}^{Q} (Q - D_i) p(D_i)$$
$$= (12000 - 8000) * 0.110 + (12000 - 10000) * 0.110$$
$$= 440 + 220 = 660$$

$$E[Retail\ profit] = 125 * 11340 + 20 * 660 - 80 * 12000 = \$470{,}700$$

$$E[Mfg\ profit] = 12000(80 - 35) - 100000 = \$440{,}000$$

You can apply similar methods for buy-back contracts, revenue-share contracts, and global optimization.

6. QUANTITATIVE INCENTIVES CONTRACT AND HUMAN NEEDS

The incentive contracts can be linked to human needs. The mission statements of most corporations nowadays emphasize social responsibilities and environmental stewardship, which are often linked to human needs. The supply-chain engineer can consider these factors in developing incentive strategies, such as the diversion rate-based transportation charges or the awards the senior design team devised in the waste management project for conference centers.

12. SUMMARY

This book develops supply-chain strategies using models in supply chains and some supply chain related models in microeconomics. A unique perspective in this book is to link the supply chain to human needs. We explore how supply-chain efficiency in financial measures impacts the supply chain efficiency to human needs. In most areas, the financial measures are aligned with human needs measures. In that case, the markets do wonders. However, as the society moves from supply capacity scarcity to supply capacity abundance, the two measures diverge in certain markets, such as in the food supply chain and underperforming healthcare systems, as pointed out in the introduction and various chapters. To conclude, we will discuss how the strategies developed in this book impact the food and healthcare supply chains.

1. THE SUCCESS OF STRATEGIES IN THE FOOD SUPPLY CHAIN

In the developed countries, we take for granted a steady supply of a wide variety of foods that only cost a small fraction of our income. Here we refer to the total supply and averages. In reality, there is poverty and food insecurity for certain populations. Throughout history, this has not been the case, because food supply faced many challenges. The supply-chain professionals, along with farmers, engineers and scientists, together achieved this feat.

1.1 Challenges in the food-supply chain

Most foods originate from farms. There are many challenges to produce farm products and to convert them to foods on our tables. Here are a few relevant to this book.

1. The pests and insects also love our foods.
2. Weeds compete to get the nutrient the farmers prepared for the farm products.
3. Farm products are subjects to diseases.
4. Farm products need a certain environment such as sun, temperature, altitude, soil, and water. Therefore, each product is suitable at certain locations. In all locations, they are susceptible to the large-scale uncertainty in Nature.
5. Farm products are seasonal. However, modern consumers demand all food types year-round.
6. Farm products are low in density. They need a large space. Combined with the environment and seasonality factors, the low density expands distance from where the products are produced from our tables, especially in modern population centers.
7. Most farm products have a short shelf-life, especially the ripe fruit, ripe vegetables, fresh meat, and fish. The objective for the food-supply chain is to supply the right food at the right location at right time in the right quantity with minimum cost. However, the short shelf-life prevents the supply chain professionals from pooling over time to buffer against the huge variability and uncertainty in supply and in demand.
8. Farm products are difficult to handle. Fresh foods are often fragile, especially when they are ripe. In many developing countries, there is significant food loss between the farms and the food-collection points. The farm products vary in shapes, sizes, and

consistencies. As a result, it is difficult to enjoy economies of scale, and to share. Farm products are subject to great uncertainty and waste.

Due to these challenges, a sufficient food supply is always a challenge to human beings.

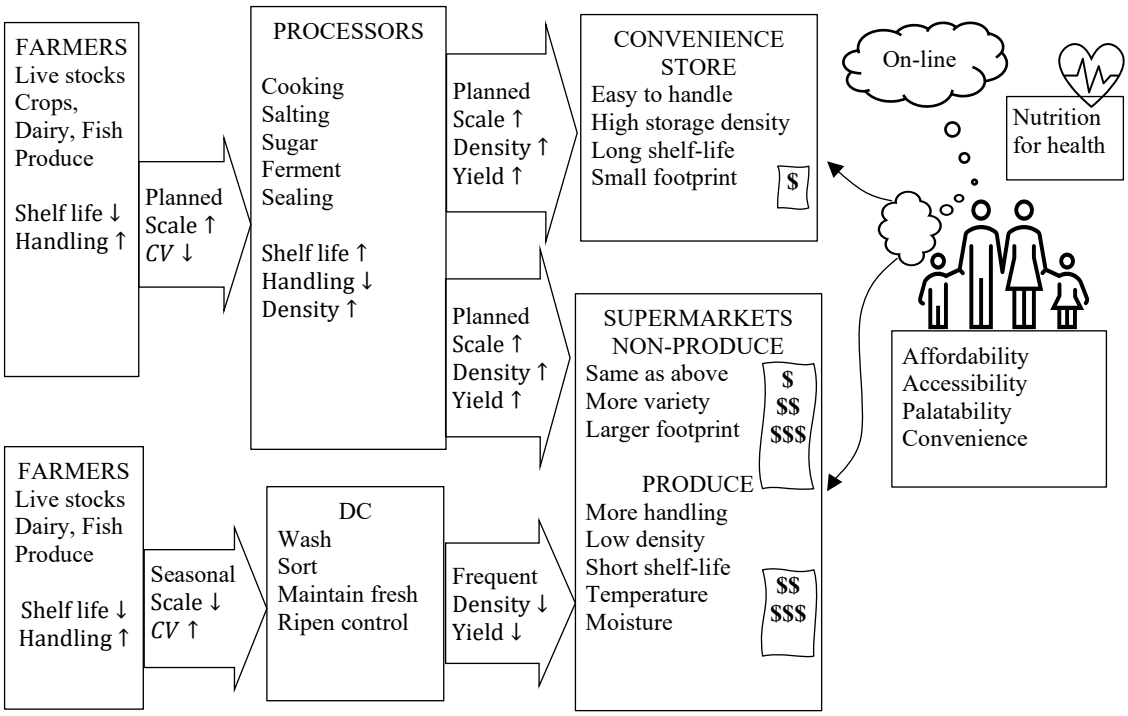


Figure 12.1. Food supply chain and food desert.

1.2 Success of the modern food-supply chain in the United States

Against all the challenges, the supply chain professionals, together with scientist and engineers, have developed a system with abundant supply of many varieties of foods, even during the pandemic. Although there is still food insecurity for some in society, the challenges are not in the total quantity, evidenced by significant food waste.

The supply chain is so efficient that people in United States spent only 12.9 percent of their income on foods, including food away from home, based on the Consumer Expenditure Survey in 2018, shown in Table 1.1 in Chapter 1. This percentage is much lower than that in the other developed countries, which range from the upper teens to over 20%.

1.3 Overview of the U.S. food-supply chain

Figure 12.1 illustrates a simplified food-supply chain for human needs. The farmers produce farm products such as crops, produce, dairy, and meat. The farmers send farm products to processors

or distributors. The food processors process the food and send them to retailers. The fresh farm products such as fruit and vegetables are also sent to distribution centers for direct distribution to the retailers without major processing.

Today, food retailing is still mostly through brick-and-mortar stores. We divide them into two major categories: convenience stores and supermarkets. The convenience stores also include gas stations, drug stores, food marts, and dollar stores. They normally carry a limited assortment of popular low-cost processed and packaged foods. The supermarkets are large stores with most food departments, and always carry fresh foods such as fruit, vegetables, and fresh meats.

1.4 Achievements in farming and the global supply chain

Pesticides and herbicides reduced the level of challenges 1, 2, and 3 greatly. The efficient global supply chain reduced the level of challenges of 4, 5, and 6 greatly. GMO seeds allow foods to grow in more and new locations with high yield and resistance to drought, flood, and diseases. The greenhouses and concentrated animal feeding operation (CAFO) also helped to reduce the challenges in the food-supply chain. The ripening technology and efficient logistics using temperature-controlled container shipping can extend the shelf- life from days to weeks. Today, the consumers on the East Coast can enjoy the fruit and vegetables from the West Coast. The consumers in North America can enjoy fruit and vegetables from South America in the opposite seasons.

1.5 Benefits of food processing with respect to the challenges

The scientist and engineers have developed many techniques using cooking, salting, drying, ripening, coloring, sealing, fermenting, and packaging, among others. They can also add preservatives. Food processing extends shelf-life from days to months. The longer shelf-life allows food to enjoy pooling in time to mitigate the cost due to variability and uncertainty.

The packaged food with uniform forms and weights is also much easier to handle, to stack, to transport, and to store. The packages can protect and help with marketing. Therefore, packaged food enjoys the benefits in economies of scale, the economies of sharing, and reducing uncertainty associated with transportation, handling, and storage.

As a result, the prices of processed fruit, vegetables, and meats are cheaper than fresh foods in the United States. The costs of fresh vegetables, fruit and meats are normally more than 20% higher than single-ingredient packaged counterpart, excluding potatoes, onions, and so on, since they have a long shelf-lives and are easy to handle.

1.6 Other benefits of food processing

Food processing adds convenience. In food processing, machines select, clear, clean, cut, mix, cook, and package food so that the foods are near read to serve on our tables. The consumers only have to get the packages out of the refrigerator or pantry, open, mix, and heat before serving. The process does not require many facilities, time, or skills. Fresh ingredients, on the other hand, require time and other resources to clear, clean, cut, mix, and cook. These processes need space, utensils, water source, drainage, a stovetop, skill, and so on.

Food processing can also add sugar, fat, seasoning, flavor, and other things to entice our appetite. Some of the artificial flavors can be made much stronger than nature can produce. In the United

States, the cost of these additives is low. They can be used to differentiate products from their competitors. Food processing can also change the appearance, smell, and the texture to make them more appealing to the consumers.

Many people, especially energy-hungry children, are naturally attracted to high-energy food that is fatty, sweet, and strongly flavored. In the United States, fats, sweets, and seasonings are low in cost. The combination of appeal and low cost allows the food processors to make processed food more palatable to the targeted population. Birth and associates showed that personal food palatability is influenced strongly by the types of food offered during a person's childhood [1], and these food preferences exhibit strong inertia. Research also shows that the human brain often favors short-term reward rather than long-term benefits, such as taste, comfort, and convenience [2]. Food suppliers can use big data and artificial intelligence to make food more palatable and to expand their market; they know how to push the taste buttons of the food shoppers. This can lead to a self-fulfilling prophecy for the food-supply chain through generations.

1.7 Supply-chain strategies and efficient food-supply chain

First, we can compare the supply chains between the convenience store group and the supermarkets. On the right-hand side in Figure 12.1, we showed two types of retailers: (1) convenience stores, gas stations, drugstores, or small food marts; and (2) supermarkets or major food outlets with at least $2 million in annual sales and containing all the major food departments. Both chains share the processed food-supply chain. The convenience stores only supply a limited assortment that is suitable in their markets: low-cost processed food that is easy to store, handle, transport, and with a long shelf-life. Some may also carry limited frozen food, also sharing the same benefits except that would incur energy cost to keep the food frozen. The major supermarket would carry the same food as the convenience stores, plus more variety and higher quality of packaged foods, frozen foods, and refrigerated foods for the more affluent customers. More importantly, the supermarket also carries a large variety of fresh fruit and vegetables, meats, and many other more difficult to carry items. We highlight the relevance to different strategies in the chapters with the numbers below.

3. Collaboration in Supply Chain(and principal-agent): Firms in both chains adopt various collaboration strategies. The only difference between the two chains is that there are certain government incentive programs that impact different chains differently. So far, the incentives are mostly financial and minimum cost.
4. Competition: Groceries are a low-margin and low-risk business with many players. Even the largest chains such as Dollar General, Walmart, and Kroger, cannot achieve much oligopoly power. Therefore, they have to compete on cost. The convenience stores limit the assortment to SKUs that are low in cost, high in sales per square foot, easier to store and handle, with a long shelf-life. They also incur fewer costs, have a smaller footprint, require less personnel (often just one person), and are easier to manage. The supermarkets carry the above and many other items, including fresh food. Fresh foods incur higher costs due to handling, low density, more maintenance, and more stringent requirements on temperature, moisture, and display. Here, the two chains compete through differentiation in assortments.
5. Pricing in Supply Chain: Same for both chains.

6. Logistics: The fresh foods incur much higher cost in transportation and storage due to temperature control, handling, low storage density, and damage. The shorter shelf-life requires shorter lead times, which can incur higher cost and lead to more cost due to shortage or overstock.
7. Economies of scale: The packaged foods enjoy more economies of scale because they are designed to enjoy the scale effect in storage, handling, transportation, and sales. higher storage density. The longer shelf life also allows longer pooling in time to enjoy scale effect.
8. Economies of Sharing and Economies of Scope: The supermarkets carry much more assortment across many departments. However, it is difficult to share resources among different departments in supermarkets. Therefore, they cannot enjoy the economies of scope.
9. Uncertainty: The long shelf-life of processed food allows the food to mitigate the cost of uncertainty by pooling over time.
10. Economies of Speed: The same for both chains.
11. Quantitative Incentive Models: The same for both chains.

2. FOOD SUPPLY AND HUMAN NEEDS

How well does the low-cost, high-efficiency food-supply chain meet human's health needs? Human health requires more than calories. Nutritional science has identified more and more micronutrients from foods that are important to human health. Protein was defined in the mid-19th century. Now we know people need vitamins, minerals, antioxidants, Omega-3 fatty acids, etc. These discoveries spawned a large nutritional supplement industry. However, the industry has yet to show strong evidence that they work as well as those found in natural food. More and more nutrients are identified by scientists even today. We do not know how many important micronutrients we have not identified that are important to our health.

Figure 1.2 in Chapter 1 showed that United States has the lowest healthy life expectancies and they are decreasing. Many have realized that our food supply does not provide the best food for our health needs. The oversupply of strong-flavored packaged food, plus the low-cost, artificially ripened fresh foods, lead to overeating of food lacking important micronutrients.

2.1 Food desert

The comparison in the last section showed that the convenient stores have a competitive advantage over the large supermarkets in low-income regions, given the advantages in cost, convenience, and palatability of packaged food. Without transportation, the people without cars do not have access to fresh food. In fact, most people in affluent regions do not have supermarkets within one mile from home because modern U.S. communities are designed based on access to cars, although the affluent population without cars can afford taxis, hire helpers, or shop online. Therefore, there is yet to be a market solution for the food desert.

2.2 Broader impact of low-cost processed food

The shortage of human nutritional needs from fruit and vegetables in United States is not limited to the food desert. A CDC *Morbidity and Mortality Weekly Report* (MMWR) showed the percentage of high school students meeting the FDA-recommended daily consumption of fruit

and vegetables in 2013 and 2017 [3]. The FDA recommends 1.5 cups of fruit for females and 2 cups for males aged 14 to 18. It recommends 2.5 cups of vegetables for females and 3 cups for males of the same age group. But in 2013, the percentages of student who consumed the recommended amounts were only 8.5% and 2.1%. In 2017, the percentages were 7.1% and 2.0%, even lower! This means that processed food is so pervasive that it impacts more people than those in the food deserts.

The National Center for Health Statistics shows the prevalence of overweight, obesity, and severe obesity among young children aged 2 to 19 from 1963 to 2018 [4]. The percentage of obese children increased from just above 4% in the 1963 to 1965 survey to 19.3% in the 2017 to 2018 survey! Severe obesity increased from 1% in the 1971 to 1974 survey to 6.1% in the 2017 to 2018 survey!

The data show that the cost- and efficiency-driven food-supply chain is not always aligned with human nutritional needs.

2.3 Organic food-supply chain

Today, many people have realized the limitations of the low-cost food-supply chain in meeting human nutritional needs, as well as social and environmental needs. There are movements in organic food, community-supported agriculture (CSA), farmer's markets, local food, slow food, and many other areas. However, they all face some challenges. We can compare organic food with the mainstream food supply chain. The Organic Trade Association reported that organic food had seen double-digit increases and reached $56 billion in sales in 2020, which is more than 5% of total food sales. Since the prices of organic foods are higher, the percentage by volume is lower. Therefore, the total sales are still much lower than the mainstream food-supply chain by volume. We can use the strategies in this book to understand why the prices are higher:

1. Organic food must face the challenges with pests and weeds. Without the chemicals, the food incurs higher cost. However, they have a smaller environmental footprint, which is external to the financial systems.
2. In addition to a higher cost in farming, organic food also loses the economies of scale in the current market.
3. The smaller scale also reduces the pooling effect and among smaller volumes. As a result, the cost of too much or too little is also higher.

As the demand continues to increase, the organic food-supply chain can enjoy more scaling, sharing, and pooling effect. The lower cost can increase the demand. A virtuous spiral can also be possible. However, there are still cost disadvantages with today's technology.

However, if the supply-chain objectives include human health, and there are ways to link health with the finances, the food-supply chain can serve human needs better.

3. HEALTH, HEALTHCARE, AND HUMAN NEEDS

As already discussed in Chapter 3, humans need health. However, the healthcare system in the Unites States is incentivized to provide care. As a care provider, the U.S. healthcare system has performed well, reflected by the most number of years if we subtract the WHO's healthy life expectancy from the life expectancy, shown in Figure 3.5. The contrasts between Figures 1.2 and 3.5 show some of the problems in the food-supply chain and the healthcare systems.

Due to the information asymmetry, lack of transparency, and necessary government involvement, the market forces and competition in healthcare are skewed. The financial incentives of healthcare providers, the insurers, the drug manufacturers, and intermediary organizations, are often misaligned with the healthcare objectives. Many doctors help the patients with health. However, the management often incentivizes them in financial terms. As a result, many doctors are frustrated. There is significant physician burnout.

Things are changing. The new pricing transparency rule already made some hospitals post their charges online. Many doctors spend time to help patients to change their lifestyle for healthy living. There are fixed-cost service providers. More and more health services are being created.

3.1 What can we do?

As supply-chain professionals, what can we do when faced with these big questions about human needs, nutritional needs, human health, and healthcare? We can learn the market, the incentives, and market forces. We can consider the human needs in addition to financial needs, and find ways to align the financial objectives and human needs objectives. We can help the suppliers, customers, colleagues, and leaders to understand the human needs beyond cost and profit. More and more people and organizations understand the importance of social and sustainability issues associated with the supply chain. We can refer to the mission statements and strategic plans to make your points. For example, the mission statement of the Georgia Institute of Technology is “to develop leaders who advance technology and improve the human condition.” The human condition includes social, environmental. and financial dimensions.

REFERENCES

Chapter 1

1. World Health Organization, "Healthy Life Expectancy (HALE) at Birth," 2000–2019, https://www.who.int/data/gho/data/indicators/indicator-details/GHO/gho-ghe-hale-healthy-life-expectancy-at-birth

2.CSCMP,https://cscmp.org/CSCMP/Academia/SCM_Definitions_and_Glossary_of_Terms/CSCMP/Educate/SCM_Definitions_and_Glossary_of_Terms.aspx?hkey=60879588-f65f-4ab5-8c4b-6878815ef921, July 12, 2019, accessed February 21, 2020.

3. U.S. Bureau of Labor Statistics, "Consumer Expenditure (CE) Surveys, Annual Calendar Year Tables, 2018, https://www.bls.gov/cex/tables.htm#annual, accessed May 1, 2020.

4. Pew Research Center, "2020 Election Reveal Two Broad Voting Coalitions Fundamentally at Odds," https://www.pewresearch.org/fact-tank/2020/11/06/2020-election-reveals-two-broad-voting-coalitions-fundamentally-at-odds/, accessed December 19, 2020.

5. Centers for Medicare & Medicaid Services, "National Health Expenditures by Type of Service and Source of Funds, CY 1960–2018," published by https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/NationalHealthAccountsHistorical, accessed May 3, 2020.

6. Paulina Cucalon, Willem Hartog, Mary Smith, Andre Vallejo, and Georgia Warnock, "Healthcare around the World," ISyE 4301 term project report, October 29, 2020.

7. Anna Wilde Mathews, Tom McGinty, and Melanie Evans, "How Much Does a C-Section Cost? At One Hospital, anywhere from $6,241 to $60,584." *Wall Street Journal*, Feb 11, 2021. https://www.wsj.com/articles/how-much-does-a-c-section-cost-at-one-hospital-anywhere-from-6-241-to-60-584-11613051137, accessed February 12, 2021.

Chapter 2

1. Porter, Michael E, "The Five Competitive Forces that Shape Strategy," *Harvard Business Review*, 2008-01, Vol.86 (1), pp.78-137.

2. Brandenburger, A., and Nalebuff, B., "Co-opetition." *Currency,* Doubleday, 1997.

3. Janice H. Hammond, "Barilla SpA (A)," Case study, revised May 25, 2008.

4. V.G. Narayanan, and Ananth Raman, "Aligning Incentives in Supply Chain." HBR Article: (HBR OnPoint Enhanced Edition), Nov. 1, 2004.

Chapter 3

1. "How does Singapore's Healthcare System Work?" https://www.pacificprime.sg/blog/singapores-healthcare-system/#:~:text=Singaporeans%20enjoy%20universal%20healthcare%20%E2%80%93%20meaning,public%20healthcare%20is%20not%20free, accessed June 9, 2021.

2. *Wall Street Journal*, "How Drug Prices Work," a video. https://www.youtube.com/watch?v=aeG2lWxYO_Y, accessed September 22, 2021.

Chapter 6

1. “How does Singapore’s Healthcare System Work?” https://www.pacificprime.sg/blog/singapores-healthcare-system/#:~:text=Singaporeans%20enjoy%20universal%20healthcare%20%E2%80%93%20meaning,public%20healthcare%20is%20not%20free, accessed June 9, 2021.

2. *Wall Street Journal*, “How Drug Prices Work,” a video. https://www.youtube.com/watch?v=aeG2lWxYO_Y, accessed September 22, 2021.

Chapter 7

1. Harris Ford, “How Many Parts to Make at Once,” *Factory*, 1913, Vol. 10, No. 2, pp 135–136. Reprinted in *Operations Research*, Vol. 38, No. 6, p. 947.

2. Gallioni, G., Montgros, X, Slagmulder, R., Wassenhove, L., and Wright, L., “Inventory-Driven Cost,” *Harvard Business Review* Case, March 2005.

Chapter 8

1. Chernev, Alexander; Böckenholt, U.; Goodman, J.K. (2015). "Choice overload: A conceptual review and meta-analysis." *Journal of Consumer Psychology*, 25 (2):333–358.

Chapter 10

1. Stanley K. Smith, Terry Sincich, “An Empirical Analysis of the Effect of Length of Forecast Horizon on Population Forecast Errors.” *Demography*, Volume 28, Issue 2, pp. 261-274, May 1991.

Chapter 12

1. Birch, L., Savage, J. S., and Ventura, A., 2007. “Influences on the Development of Children's Eating Behaviours: From Infancy to Adolescence,” *Canadian Journal of Dietetic Practice and Research*; Vol. 68, No. 1 (Spring 2007): S1-S4,S6.

2. Scaglioni, S., De Cosmi, Ciappolino, V., Parazzini, F., Brambilla P., and Agostoni, C., 2018. “Factors Influencing Children’s Eating Behaviours,” *Nutrients*, 10 (6), p. 706, 2018.

3. Lange S.J., Moore L.V., Harris D.M., et al., “Percentage of Adolescents Meeting Federal Fruit and Vegetable Intake Recommendations—Youth Risk Behavior Surveillance System, United States, 2017.” MMWR *Morb Mortal Wkly Rep* 2021;70:69–74, https://www.cdc.gov/mmwr/volumes/70/wr/mm7003a1.htm?s_cid=mm7003a1_w#T2_down.

4. Fryar C.D., Carroll M.D., Afful J. “Prevalence of Overweight, Obesity, and Severe Obesity Among Children and Adolescents Aged 2–19 years: United States, 1963–1965 through 2017–2018. NCHS Health E-Stats, 2020, https://www.cdc.gov/nchs/data/hestat/obesity-child-17-18/obesity-child.htm#Suggested%20citation, accessed March 3, 2021.

5. Maggie McNeil, “U.S. Organic Sales Soar to New High of Nearly $62 Billion in 2020,” Organic Trade Association, https://ota.com/news/press-releases/21755, accessed July 5, 2021.