Title
Decentralized Control of Stochastic Dynamic Systems with Applications to Resource Allocation and Portfolio Management

Author
Cai, Huaning

Publication Date
2012

Peer reviewed|Thesis/dissertation

Decentralized Control of Stochastic Dynamic Systems with Applications to Resource Allocation and Portfolio Management

by

Huaning Cai

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Engineering - Industrial Engineering & Operations Research

in the

GRADUATE DIVISION
of the
UNIVERSITY OF CALIFORNIA, BERKELEY

Committee in charge:
Professor Andrew E.B. Lim, Chair
Professor Zuo-Jun Shen
Professor Pravin Varaiya

Fall 2012

Decentralized Control of Stochastic Dynamic Systems with Applications to Resource Allocation and Portfolio Management

Abstract

Decentralized Control of Stochastic Dynamic Systems with Applications to Resource Allocation and Portfolio Management

by

Huaning Cai

Doctor of Philosophy in Engineering - Industrial Engineering & Operations Research

University of California, Berkeley

Professor Andrew E.B. Lim, Chair

Modern engineering and social systems are often too complex to be managed by a centralized agent. Instead, such systems are commonly structured with multiple decentralized agents each responsible for managing a subset of the system, but the resulting system performance depends on the aggregate of the decisions made by decentralized agents. Local agents' decision makings often exhibit selfish behavior as they seek to optimize their own objectives under their localized models, which if left uncoordinated can lead to substantial loss of efficiency compared with the system that can be optimized by a single (hypothetical) centralized agent. In this dissertation, we seek to study the fundamental issues of how to efficiently manage large-scale and multi-agent stochastic dynamic systems, especially on how to device efficient coordination mechanisms that would optimize system performance under various constraints that are unique to decentralized systems.

In the first part of this dissertation we study decentralized control of a general class of stochastic dynamic resource allocation problems that have many applications. We consider a stochastic system in which multiple decentralized agents allocate shared system resources in response to customer requests that arrive stochastically over time. Each agent is responsible for a subset of the allocation decisions which it makes according to a dynamic allocation policy obtained by maximizing his own expected profit subject to a potentially mis-specified model of the way in which shared resources are consumed by other agents. We introduce the notion of a transfer contract which specifies how agents compensate one another whenever resources are consumed and establish the existence of contracts under which the decentralized system has no efficiency loss relative to centralized optimality. We also show that this property is insensitive to mis-specification by each agent of the dynamics of resource consumption by others in the system. An explicit characterization of the optimal transfer contract and an iterative decentralized algorithm for computing it is also provided. In the language of duality, contracts are analogous to shadow prices and the iterative algorithm has the favor of a dual update method, but strong duality and convergence of the it-

erative algorithm to the set of optimal contracts are guaranteed without assumptions of convexity.

In the second part of this dissertation we study a class of related decentralized control problems but specialize to portfolio and risk management. Many financial institutions typically trade in multiple correlated markets. While centralized portfolio optimization over all trading decisions is ideal, it is generally not possible due to the complexity of each market, and firms typically adopt a decentralized setup in which trading in each market the responsibility of a particular desk. Decentralized portfolio optimization, however, is complicated by the fact that different agents are commonly only well informed about their own investment universe (proprietary research and forecasts, etc) and prefer to keep this private, and have their own incentives which they optimize on the basis of their limited models. It is well known, however, that the aggregate performance of such a system can be extremely inefficient due to the loss of diversification. In this dissertation, we formulate a multi-agent dynamic portfolio choice problem and study how to improve its efficiency. We show that an internal system of swap contracts, which define internal cash transfers between agents, can be used to facilitate risk sharing and induce agents to choose portfolios that as a collection are optimal for the firm. Conceptually using swap contracts is similar to performance benchmarking that is often employed in the finance literature for decentralized portfolio management, but our new approach offers a significant advantage in that the swap contracts can be constructed in decentralized manner without requiring an all-knowing central agent. We provide an explicit characterization of the optimal swap contracts and an iterative algorithm for computing them that can be implemented without compromising proprietary agent level data.

Throughout this dissertation, we also discuss various important issues surrounding decentralized control of stochastic dynamic systems, including but not limited to approximation methods, performance attribution, sensitivity analysis, and fairness issues, etc.

In memory of my mother, Weiyin Wang, and to my father, Liangrong Cai, with love and eternal appreciation.

Contents

List of Figures

List of Tables

Acknowledgments

First and foremost, it is difficult to overstate my gratitude to my advisor, Professor Andrew Lim. Andrew has been a constant source of great ideas and inspiration. He has shown tremendous enthusiasm in my research. He is always there asking me tough questions, providing me with constructive feedback, and helping me overcome obstacles one after another. I could not have completed this dissertation without his incredible help and support. Moreover, Andrew has made great efforts in securing financial supports for my entire graduate study, and he is always making sure that I have been taken good care of and can fully concentrate on my study and research. I cannot ask for a better advisor.

I would like to thank Professor Zuo-Jun Shen and Professor Pravin Varaiya for serving on my dissertation committee. I would like to thank them for their tough questions and insightful comments that have made this dissertation a better one. I would also like to thank Professor Ying-Ju Chen and Professor Xuanming Su for serving on my qualification committee and their constructive feedback in shaping my early research.

I would like to thank many of my fellow students for providing me with a stimulating and fun environment in which to learn and grow. I am especially grateful to Ankit Jain, Lian Yu, Poomyos Wimonkittiwat, Gah-Yi Vahn, Shea Chen from my research group, and Tianhu Deng, Yong Liang, Dawei Chang, Xingwei Wu, Yao Ma, Mengshi Lu, Nguyen Le Truong, and many others who have made my graduate life at Berkeley a much more interesting and enjoyable one.

I am greatly indebted to my late mother, my father, and my elder-sister for their unconditional love and support. They are always there when I needed encouragement and they always inspire me to become a better person. I also owe a great deal to my parents-in-law for loving me like their own son, and they have been nothing but supportive in my doctorate pursuit. Finally, I am extremely grateful to my loving wife, Dr. Yuxiong He. Finishing a doctorate is a long and incredibly strenuous process, with loads of ups and downs. I would not have accomplished it without having her by my side. She is forever the first one to cheer me up, tell me I am the best, and that I can achieve anything. I dedicate this dissertation to her.

Chapter 1

Introduction

1.1 Motivation

Optimal control of stochastic dynamic systems have long been the central focus of the operations research and management science community. To date, much of the research efforts have taken a centralized perspective, assuming there exists an centralized agent, who oversees the entire system, and is able to collect all necessary data and information, to build a complete model, and optimizes all decisions for the entire system. However, many modern engineering and social systems are too complex to be managed in such a manner. More commonly, such systems are populated with multiple decentralized agents, each responsible for a subset of the system, specializing in different and complex subtasks, and typically only interested in optimizing their own local objectives that are often not aligned with that of the system. The resulting system performance depends on the collection of the agent level decisions and the way they interact with one another. In short, centralized decision making rarely occurs, and a decentralized control perspective is more consistent with what is being done in practice.

The main research goal of this dissertation is to study how to optimize the performance of complex stochastic dynamic systems that are managed by multiple agents. Decentralized control is an essential when (i) a system is too complex to be controlled in a centralized manner, (ii) it is populated with multiple agents who perform specialized and complex tasks, (iii) agents are selfish and are interested in maximizing their own objectives which are often not aligned with that of the system, or (iv) when relevant data or information is distributed amongst multiple agents and can not be aggregated (as required when formulating a centralized stochastic optimization problem) due to privacy issues or the complexity of the aggregating task. Next we discuss several real-world examples involving complex stochastic dynamic systems where needs for decentralized control naturally arise. Through these examples we highlight key questions that we aim to investigate in this dissertation.

1.2 Examples

Example 1.2.1 (Airline alliance revenue management).
Since the deregulation of US domestic market in 1978, most US airlines have formed alliance with international carriers to share flights, frequent flyer programs, ground operation, and etc. In an airline alliance where multiple partner airlines share common inventory/seats, prices are still typically set in a decentralized manner by each of the participating airlines. While system-wide profits can be improved (in theory) by jointly optimizing over all prices offered by all participating airlines, centralized pricing in an alliance is never done. One reason is that such a joint optimization problem would be enormous; a second is that the task of integrating demand models and the IT systems from each of the airlines into a single system would be overwhelming. A third and more fundamental reason is that demand forecasts and data for the customers of each of the airlines is typically proprietary and airlines are often reluctant to share this information with their partners in the alliance or with a centralized decision maker, so the information required to make optimal centralized decisions can not be collected in one place. The main challenge is to design incentive structures so that optimal decentralized pricing leads to good performance of the system.

Example 1.2.2 (Optimal pricing and inventory control).
Large firms are typically structured with multiple business units (e.g. one responsible for pricing and marketing, the other responsible for manufacturing, another responsible for logistics and distribution, and etc) that make decisions dynamically over time in a uncertain environment. Efficient running of such firms requires effective coordination among these units. One typical approach that has been well studied in the OR/MS literature is to model an aggregate system and to jointly optimize over decisions of all business units: e.g. joint pricing and inventory decisions. Though interesting and undoubtedly useful in some contexts, the practice of jointly optimizing over the dynamic decisions of different business units is difficult for a number of reasons: (i) the tasks performed by each group are complex and require specialized domain knowledge; (ii) it is difficult to integrate data, information, and IT systems from the different groups; (iii) groups may be in different countries and speak different languages. Instead, we typically see each group optimizing its own performance measures. The challenge in system design is once again to devise proper incentive structures so that optimal control policies for the different business units are also optimal for the firm.

Example 1.2.3 (Portfolio and risk management).
Financial institutions typically trade in multiple correlated markets. One traditional approach to portfolio and risk management is to optimize over all trading decisions centrally. However, this is generally not possible due to the uniqueness and complexity of each markets, and firms typically hire multiple specialized agents to manage different markets (e.g. equity, fixed income, commodity, etc). While the firm is concerned

with the risk-return profile generated by the trading decisions selected firm wide, individual agents are often only well-informed about their own market, and optimize their local risk-return profiles. The trading decisions they make could be far from optimal given trading decisions from other agents and could lead to over-exposure of the firm to certain risk factors. The goal is to design incentives for each agents so that firm wide optimal trading decisions are chosen.

From these examples we can see that the needs for decentralized control arise naturally. In this dissertation we seek to understand how decentralized stochastic dynamic systems can be structured and incentives can be set up so that decentralized agents will make decisions that are good for the system as a whole. More specifically, we divide this dissertation into two main components: in the first part we study a general class of decentralized stochastic dynamic resource allocations problems, which has many real-world applications, such as the above mentioned optimal pricing and inventory control, airline alliance revenue management, and etc; in the second part we study a class of related problems but specialize to portfolio and risk management, where the kind of coordination mechanisms required are quite distinctive from the ones used for the resource allocation problems. We next provide brief introductions and literature reviews for these two main components of this dissertation.

1.3 Decentralized Stochastic Dynamic Resource Allocation

Consider a system made up of multiple decentralized agents sharing a limited pool of resources. Each agent offers a menu of products (each of which is a bundle of resources) and receives requests for these items stochastically over time. On receiving a request, an agent needs to decide whether it should be accepted or rejected. If rejected, the customer departs; if accepted, the bundle is assembled using resources from the common pool and delivered to the customer in return for a pre-defined payment. Consumption of shared resources also comes at a cost which is paid by the agent accepting the request as compensation to other agents in the system. The cost of resources and the manner in which it is divided amongst the other agents in the system is defined by a *transfer contract*. Every agent earns a profit which is the sum of income from his own customers whenever a request is accepted and transfer payments from other agents (whenever they accept a request from one of their customers) minus the cost of resources used to manufacture allocations to his own customers. For a given set of transfer contracts, each agent solves a decentralized stochastic control problem in which he maximizes expected profit subject to his model for the arrival of customer requests and a possibly mis-specified model for the way in which resources are allocated by other agents. The solution of this problem determines an allocation policy for each agent. Clearly, the behavior of each agent (and hence the aggregate system) is a function of the chosen transfer contracts as well as the model adopted by each agent for the evolution of the system. The goal of the first part of this dissertation is to understand how these contracts should be chosen so as to maximize the net revenue of the overall system. One example of our system is an airline alliance where multiple airlines (agents) share fight legs (resources) which are bundled and sold as itineraries (products) to customers that arrive stochastically over time.

The problem we have just formulated is a decentralized stochastic dynamic resource allocation problem. We note, however, that while resource allocation has a long history (see for example Arrow et al. [21]) there are relatively few papers dealing with this problem in a decentralized stochastic dynamic setting. One related line of work is the literature on decentralized control dating back to the papers of Marschak [25], Arrow and Hurwicz [3], and Radner [28] in the economics literature, and Witsenhausen [42] and Ho [18] in the systems and control literature (see also [22, 29, 30, 31, 32, 43]). In these papers, agents are cooperative but each sees a different function of the aggregate system state, and the primary concern is describing the aggregate system performance when agent policies are constrained to be functions of their observations. In contrast, our agents are *self-interested* and maximize their own profits subject to a private but mis-specified model of the way in which system resources are consumed. Specifically, although each agent is assumed to accurately know the probability of receiving a request for a resource bundle in his menu, his model for resource consumption by other agents may be mis-specified. Quite surprisingly, we show that transfer contracts can be found under which there is no efficiency loss relative to an optimal fully knowledgeable centralized agent, *even if agent level models are mis-specified* along the lines we have described.

While our analysis of the decentralized allocation problem proves existence of coordinating transfer contracts and provides an explicit construction, one limitation of this result is that the contract depends on the value function of the centralized problem (which is not surprising since it does enable the decentralized system to achieve centralized efficiency). This is problematic because our study of the decentralized system was motivated by the view that agents with the requisite knowledge to solve the centralized problem simply can not be found in many applications, so stopping at a "solution" that can only be computed by such an agent would be disquieting, to say the least. With this in mind, we address also the problem of computing the optimal set of contracts in a decentralized manner. We propose an iterative algorithm that only requires decentralized agents to update and exchange their valuations of shared resources, which they compute using their private mis-specified models, and show that it converges to the optimal transfer contract. More generally, transfer contracts can be interpreted as shadow prices in a stochastic dynamic resource allocation problem, and the iterative procedure as a dual update algorithm, so it is interesting that existence of optimal contracts (strong duality) and convergence of the update algorithm can be guaranteed without any assumptions about convexity.

For other related work, decentralized decision making has attracted a lot of attention in the OR/MS literature, but the primary focus has been on single period problems. See for example the survey of Cachon [11] and the edited volume by Simchi-Levi et al. [34]. A key distinction of our study is that we focuses on decentralize control of stochastic and dynamic problems, an issue which has received substantially less attention in the OR/MS literature. Two exceptions are recent papers by Bertsekas [6] and Adelman and Mersereau [2]. These papers consider weakly coupled stochastic dynamic optimization problems in which control variables across agents are coupled by convex separable constraints though state variables for different agents are otherwise independent. The aggregate system is decoupled by dualizing the control variables in the dynamic programming equations. A key distinction in our problem is that the agents in our problem share the same state variable and hence are strongly coupled.

Another closely related work is a recent paper by Moallemi and van Roy [27]. They study resource allocation in the message passing context, where message-based incentives similar to our transfer contracts also generalize the notion of Lagrangian multipliers by allowing them to vary across agents and resource consumption levels. They describe a distributed and asynchronous message-passing algorithm for computing equilibrium messages and allocations, and demonstrate its merits in the context of network resource allocations problem. Though very similar in spirit, their work differ from ours in two aspects. First it is restricted to static setting; second they require strong assumptions such as convexity to establish optimality and convergence results. Our work is more general as we study stochastic dynamic problems, and we do not require convexity in establishing convergence of our iterative (i.e. dual-update or message-passing) algorithms.

It is important that we discuss the literature of airline alliance revenue management, which we mention earlier and will be used again as an application to illustrate our methodologies. Again much of the studies from the literature have taken a centralized perspective (see Talluri and van Ryzin [36] for an extensive survey), but the industry trends are clearly moving towards more cooperation and alliances formation [40]. There has been very limited work done for airline alliance, and there are only two papers that we are aware of by Boyd [9] and Wright et al. [44]. We note however, that Boyd's paper [9] is for deterministic problems while the focus of Wrigth et al. [44] is evaluating the performance of particular transfer contracts, and does not address the issue of whether it is possible to coordinate a more general system with transfer contracts, nor the impact of model mis-specification by agents in the system (every agent is assumed to be perfectly informed about other agents). In a more recent paper, Hu et al. [20] proposes a two-stage hierarchical game-theory approach. In the first stage game, airlines negotiate for fixed proration rates as a mean to share revenue. In the second stage game, airline operate independent inventory systems to maximize their own expected revenue. Though their analysis is interesting and clearly has merits, they rely on rather simplified static approximation of the dynamic problem, and hence remains largely restricted to the traditional single-period framework.

As a final note, we discuss the paper by Li, Lim and Shanthikumar [24] which studies the problem of decentralized control of an $M/M/1$ queue (which obviously involves only two control agents, one for the arrival rate and the other the service rate). Like the present work transfer contracts that coordinate the system are also constructed in [24] and a decentralized method for computing them is also proposed. There are several key differences, however. Firstly, the present work is set in discrete time whereas [24] is continuous. Secondly, the present work involves systems with more than two decentralized agents. Thirdly, the stochastic system in the present work applies to a larger class of applications with more complex state spaces, whereas [24] concerns a particular two-agent single server queue with a one-dimensional state space. Fourthly, the convergence analysis in [24] exploits the continuous time structure of the problem it studies whereas the analysis in the present work, aside from being shorter, also develops techniques that are more easily extended to other discrete as well as continuous time systems.

1.4 Decentralized Portfolio and Risk management

Consider a financial firm that is interested in trading in multiple correlated financial markets. Centralized portfolio and risk management is ideal, but often the aggregate system is too complex to be managed that way, and multiple agents (portfolio managers) are hired to manage the decisions in different markets (e.g. equity, fixed income, commodity, real estate, etc). Each decentralized agent solves a stochastic optimal control problem in which he maximizes his own risk-adjusted expected return conditional on a private stochastic model of the market in which he is investing in. The firm's net positions are the collection of these holdings but the firm is interested in the risk-adjusted return of the aggregate investment. While decentralized investment management is necessary and the cost to the firm of agents not coordinating can be substantial, little is understood as to how this can be done. In this dissertation, we formulate this problem as a decentralized stochastic optimal control problem and study a mechanism, based on the idea of risk transfer, for optimizing system efficiency.

This dissertation has two main contributions. Firstly, we show that it is possible to coordinate locally knowledgeable decentralized agents by introducing a system of swap contracts that define internal cash transfers between them. Intuitively, these contracts facilitate risk transfer and set a price for risk, which allows agents to be rewarded for reducing the firm's exposure to a particular risk source and penalized for taking it over a desirable level. As such, it is a mechanism for aligning the incentives of the agents and the aggregate system. Secondly, we introduce an iterative approach for constructing optimal swap contracts that can be implemented without any of the agents needing to reveal private model information to other agents or the centralized system. The value of this algorithm is that is allows optimal contracts to be constructed without a centralized agent with knowledge of the integrated system model.

The importance of coordination in decentralized (or delegated) portfolio management has been recognized for many years. In a seminal paper, Sharpe [33] considers the portfolio choice problem with multiple portfolio managers (agents) who do not want to share their private information with the central manager. By assuming short sales are allowed, all parties have a consensus on the covariance matrix, and all agents follow the same set of securities, Sharpe [33] develops an instruction for each agent so that he will choose portfolios that are globally optimal for the firm. However, no exact solution is available when agents follow non-overlapping securities. This open question is solved by Elton and Gruber [15], who show that global optimum can be achieved by requiring those portfolio managers to follow a specifically designed portfolio rule. In both Sharpe [33] and Elton and Gruber [15], the rules designed by the authors do not require the agents to reveal their private information (e.g. return forecasts). However, a major drawback of their approaches is that they ignore the incentives of the agents and disallow them to maximize their personal earnings. In other words, it will be difficult to enforce such rules in practice. More recently, Binsbergen et al. [37, 38] propose a different approach that utilize performance benchmarks to align incentives between decentralized agents and the central manager. Using a relatively simple framework with two agents managing non-overlapping securities, and a strong assumption of the existence of an all-knowing central manager, Binsbergen et al. [37, 38] shows that there exist stochastic benchmarks that could induce local agents to choose portfolio holdings that in the aggregate are optimal for the firm. While benchmarks are closely related to the cash transfers studied in this dissertation, the key difference is how they are constructed. In Binsbergen et al. [37, 38], an central agent with knowledge of the dependence structure between all assets in all markets - and hence capable of constructing the optimal contract by solving the centralized problem - is assumed to be available, which is problematic since the absence of such an agents is a major motivation for decentralization.

Similar to the decentralized stochastic dynamic resource allocation problems, this work is also closely related to the literature on decentralized control (see for example [17, 18, 19, 22, 29, 31, 39, 42]). An essential difference is that agents in these papers are cooperative and all know the dynamics of the aggregate system, but are constrained to policies that are functions of their observations (e.g. a subset of the aggregate system state). In contrast, our agents are interested in optimizing their own objectives and have partial knowledge of the system dynamics, and the goal is to provide incentives through risk transfer such that the collection of agent-level optimal policies is optimal for the aggregate system (see [12, 13, 24] for other applications).

Finally, another line of closely related work is the optimal risk-transfer literature (Artzner et al. [4], Barrieu and Karoui [5], Li et al. [23]), where the objective is to minimize system risk according to some risk measure by transferring risk between different parts of the system. The focus has been to characterize the structure of the optimal risk transfer using different notions of convex risk measures, and they generally take a centralized approach. Our work is different in two major aspects, first we take on a decentralized perspective, second our goal is to achieve optimal control of the system while the risk-transfer literature doesn't explicitly consider optimal control in the presence of transfer contracts.

1.5 Organization of this Dissertation

The remaining of the this dissertation is organized as follows.

- In Chapter 2 we study the decentralized stochastic dynamic resource allocation problems. We characterize optimal revenue transfer contracts that can be used to coordinate local agents' pricing decisions to achieve central optimality. We show that this property is robust to possible mis-specification by each agent of the dynamics of resource consumptions by other agents in the system. We further provide an iterative algorithm for computing the sharing contracts in decentralized manner without requiring agents to reveal their private informations, and we prove convergence.

- In Chapter 3, we study the issue of computational efficiency of our proposed coordination mechanism in Chapter 2. We propose two different approximation methods that are more scalable for practical implementation. The first approach utilizes static linear programming approximation, and by method of inventory splitting, we show that the resource allocation decisions can be decentralized efficiently. We further adapt our iterative algorithm and show that at equilibrium local agents' pricing policies will also be optimal for the central agent within the static approximation framework. The second approach utilizes dynamic affine value function approximation, where we restrict the search of optimal contracts to affine functions. We show that approximate dynamic contracts can be computed efficiently within our proposed decentralized framework.

- In Chapter 4, we study the decentralized portfolio and risk management problems. We devise a internal system of swap contracts, which defines internal cash transfer between agents that facilitates risk sharing and can induce them to choose portfolio positions that in aggregate will be optimal for the firm. We further provide an iterative algorithm that can be used to construct the optimal swap contracts without compromising agent level proprietary data, and we prove convergence.

- In Chapter 5, we discuss several important issues concerning efficient decentralized portfolio and risk management. First we propose a risk attribution method based on principal component analysis, which offers a way for the firm to identify risk factors that contributes the most to the firm's aggregate risk, and hence swap contracts can be prioritized to significantly improve system performance even with only partial implementation. Second we discuss the sensitivity of the swap contracts' efficiency gain to model assumptions, and offer strong arguments supporting the use of swap contracts, especially in the presence of model uncertainty. Finally, we address the issue of fair allocation of the surplus utility generated as a result of the efficient coordination among agents, and show that our swap contracts can be extended easily to ensure agents are provided with sufficient incentive to participate.

- In Chapter 6, we conclude this dissertation and discuss future research opportunities.

Chapter 2

Decentralized Stochastic Dynamic Resource Allocation

Many problems can be considered as multi-agent stochastic dynamic resource allocation problems where needs for decentralization arise naturally. Examples include service pricing and control in an outsourced environment, revenue management for an airline alliance, smart-grid operation, and etc. In this chapter we formulate a general class of decentralized stochastic dynamic resource allocation problems. We consider a system managed by multiple decentralized agents. Each agent manage a subset of the products (bundles of raw resources) locally but interact with other agents at the system level where a limiting pool of resources are shared. We consider stochastic demand where requests for products arrive overtime according to locally poisson processes. On receiving a request, the selling agent can decide to accept the request, thus earning a revenue and consumes resource from the system; or reject the request, and the customer departs the system. Local agents typically seek to maximize their own revenue without worrying about the cost to the system, therefore if left uncoordinated can lead to substantial efficiency loss for the system. In light of this typical selfish behavior exhibited by decentralized agents, we introduce the notion of transfer contract to coordinate decision making of the decentralized agents. Transfer contract specifies the amount an agent need to pay to other agents when a request is received and accepted, which serves as compensation to other agents for their future opportunity cost of the resources that are being consumed. We investigate in this chapter how transfer contract should be set up so as to maximize the net revenue of the overall system.

This chapter is organized as follows. Section 2.1 describes the models for resource, products, product demands, and how revenue is generated from allocation decisions. Section 2.2 and 2.3 formulate a centralized and decentralized approach for coming up with allocation decisions for the system respectively. Section 2.4 characterizes the optimal transfer contracts, and establish results about existence and optimality. Section 2.5 provides an iterative algorithm for constructing the optimal contract in decentralized manner respecting information sharing constraints and non-existence of an all-knowing central agent, as well as a case study of airline alliance revenue management. Section 2.6 proves convergence of the iterative algorithm. Finally Section 2.7 concludes this chapter.

2.1 General Description

In this section, we describe the overall system, the models for resources, products (resource bundles), product requests, and how revenues are generated from allocation decisions. In subsequent sections, we formulate both a centralized as well as decentralized approach for coming up with allocation decisions for the system described in this section. In the centralized model, allocation decisions are optimal policies for a single agent who optimizes system revenue under complete knowledge of the demand statistics for all products. In the decentralized model, allocation decisions for different products are made by different agents, where each agent only knows the demand statistics for the products he is managing, and makes allocation decisions by optimizing profits from his own sales.

Our resource allocation problem can be viewed as a make-to-order system that offers a menu $\mathcal{J} = \{1, \cdots, n\}$ of n different products. We denote by $\mathcal{L} = [\mathcal{L}_1, \cdots, \mathcal{L}_K]' \in \mathbb{R}^K$ ($\mathcal{L}_i \geq 0$) the initial inventory level of the K different resources stocked in the system. The system receives requests for products in $\mathcal{J}$ stochastically over time. We write $\widetilde{d}_{jt} = 1$ if a request for product j is received at time t and $\widetilde{d}_{jt} = 0$ otherwise, and denote by

$$q_{jt} = \mathbb{P}(\widetilde{d}_{jt} = 1, \widetilde{d}_{j't} = 0, \forall j' \neq j), \forall j \in \mathcal{J}, \tag{2.1}$$

the probability that a request for product j is received at time t. Observe in (2.1) that at most one product can be requested during each time period, so there are $n+1$ possible events (including the possibility of no arrival) and $\sum_{j \in \mathcal{J}} q_{jt} \leq 1$. On receiving a request, a decision needs to be made as to whether it should be accepted or rejected. As previously noted, this decision is made by a single agent in the centralized case, and shared amongst multiple agents in the decentralized case, where the centralized agent is knowledgeable about all system statistics and optimizes system revenue, whereas decentralized agents optimize their own profits subject to partial knowledge of the demand statistics. We denote allocation policies by $\mu_t(x) = [\mu_{1t}(x), \cdots, \mu_{nt}(x)]$, where $\mu_{jt}(x) = 1$ if a request for bundle j at time t when inventory $x_t = x$ is to be accepted and $\mu_{jt}(x) = 0$ otherwise, and $x_t = [x_{1t}, \cdots, x_{Kt}]'$ denotes the quantity of resources available at the start of time t. We denote by $A_j = [a_{1j}, \cdots, a_{Kj}]'$ ($a_{lj} \geq 0$) the resource requirement for product j and impose the requirement that product j can only be manufactured at time t if $A_j \leq x_t$, which constrains the allocation policy

$\mu_t(x)$ to the set

$$\mathcal{U}(x) = \Big\{[\mu_{1t}(x), \cdots, \mu_{nt}(x)] : \mathcal{H} \to \{0,1\}^n \;\Big|\; A_j \mu_{jt}(x) \le x\Big\},$$

where $\mathcal{H} \subset [0, \mathcal{L}_1] \times \cdots \times [0, \mathcal{L}_K]$ is the set of possible inventory levels. The system revenue under an allocation policy $\mu_t(x)$ is

$$\begin{cases} \mathbb{E}\Big\{\sum_{t=1}^{T}\sum_{j\in\mathcal{J}} r_j \mu_{jt}(x_t)\widetilde{d}_{jt}\Big\} \\ \text{subject to:} \\ x_{t+1} = x_t - \sum_{j\in\mathcal{J}} A_j \mu_{jt}(x)\widetilde{d}_{jt},\; x_0 = \mathcal{L}, \\ \mu_t(x_t) \in \mathcal{U}(x_t). \end{cases} \tag{2.2}$$

The key difference between a centralized and a decentralized system is the way in which the allocation policy $\mu_t(x)$ is obtained. In the centralized case, the policy is obtained by a single agent optimizing the objective in (2.2) with complete knowledge of request probabilities and revenues for each product. In the decentralized case, allocation policies for different products (i.e. different components of $\mu_t(x)$) are computed by different agents maximizing their own profits subject to mis-specified private models.

2.2 Centralized Control

We now formulate a centralized version of the resource allocation problem (2.2) and characterize the value function and optimal policy. The results in this section serve as a benchmark for the decentralized model that we introduce in Section 2.3.

The centralized decision maker's objective is to maximize expected revenue subject to resource constraints (see also [36]). He does so with complete knowledge of request probabilities and rewards (q_{jt}, r_j) for each product bundle $j \in \mathcal{J}$:

$$(C)\begin{cases} \max_{\mu(\cdot)} \mathbb{E}\Big\{\sum_{t=1}^{T}\sum_{j\in\mathcal{J}} r_j \mu_{jt}(x_t)\widetilde{d}_{jt}\Big\} \\ \text{subject to:} \\ x_{t+1} = x_t - \sum_{j\in\mathcal{J}} A_j \mu_{jt}(x)\widetilde{d}_{jt},\; x_0 = \mathcal{L}, \\ \mu_{jt}(x_t) \in \mathcal{U}(x_t). \end{cases} \tag{2.3}$$

The dynamic programming equation for (C) is

$$\begin{cases} V(t,x) = \max_{\mu} \mathbb{E}\Big\{ \sum_{j\in\mathcal{J}} r_j \mu_{jt} \widetilde{d}_{jt} + V(t+1, x_{t+1}) \,\Big|\, x_t = x \Big\}, \\ V(T+1, x) = 0. \end{cases}$$

Defining $\Delta V(t+1, x, A_j) \triangleq V(t+1,x) - V(t+1, x-A_j)$, the dynamic programming equation for (2.3) can be written as [36]

$$\begin{cases} V(t,x) = \max_{\mu} \sum_{j\in\mathcal{J}} q_{jt}\mu_{jt}\big(r_j - \Delta V(t+1,x,A_j)\big) + V(t+1,x), \\ V(T+1,x) = 0 \end{cases} \tag{2.4}$$

and the optimal centralized policy is

$$\mu^*_{jt}(x) = \begin{cases} 1 & \text{if } r_j \geq \Delta V(t+1,x,A_j), \text{and } x \geq A_j \\ 0 & \text{otherwise.} \end{cases} \tag{2.5}$$

In particular it is optimal for the centralized agent to accept a request for bundle j if the generated revenue r_j is no less than the value $\Delta V(t+1, x, A_j)$ of the resources being consumed and there is sufficient inventory required to satisfy the request.

In formulating problem (2.3), it is implicitly assumed that the centralized decision maker knows all relevant system parameters including the demand probability for each bundle. In many applications, such a knowledgeable decision makes does not exist and it is more common that there are many agents where each manages a subset of the products/bundles, is knowledgable about the arrival statistics of products in his own menu but not those of others, and is interested in maximizing his own revenue. In such situations, it is not possible to formulate the centralized problem (2.3).

2.3 Decentralized Control

Several important elements that distinguish the decentralized version of the resource allocation problem from the centralized version (2.3). Firstly, *the decentralized system consists of multiple agents where each is responsible for the accept/reject decisions for a subset of bundles from the pool $\mathcal{J}$.* Decentralized agents are tied together, however, because shared resources are consumed whenever any agent makes an allocation. Secondly, *every agent is better informed about the subset of products he is managing than those that are managed by others.* Specifically, while each agent is assumed to know the probability that each bundle in his menu is requested, we do not assume that he knows the probability that shared resources are consumed by other agents, or even the number of other agents in the system. Consequently, while each agent may attempt to model the consumption of shared inventory, we only assume that the part representing requests by his own customers and his associated allocation decisions is correct, and that the component associated with resource consumption by other agents is mis-specified. Finally, *each agent maximizes his own objective function conditional on his possibly mis-specified stochastic model for resource consumption by other agents*. Shared system resources are depleted stochastically over time under the decentralized set of allocation policies, which typically differ from the optimal allocation policy of an all-knowing centralized agent.

In this section, we formulate agent level dynamics and agent level objectives. In defining the objective function for each agent, we introduce the notion of transfer contracts which define revenue sharing between agents in the system whenever resources are consumed. Intuitively, transfer contracts define a price for using shared inventory whenever an agent accepts a request which is used to compensate other agents for the loss of resources. We study the impact of transfer contracts and model mis-specification on efficiency loss of the decentralized system in Section 2.4.

2.3.1 Stochastic Model for Agent $i \in \mathcal{I}$

The following is assumed about Agent i:

- Agent i makes accept/reject decisions for a subset of bundles $\mathcal{J}_i \subseteq \mathcal{J}$, and we denote by $\mathcal{J}_{-i}$ the set of bundles managed by other agents;
- All agents can observe the remaining inventory x_t at time t (e.g. through a database);
- Inventory is depleted whenever any agent receives and accepts a request;
- Agent i makes accept/reject decisions for items in his/her portfolio $\mathcal{J}_i$. These decisions are described by a policy $\mathbf{u}_i = \{\mu_{jt}^{(i)}(x),\, j \in \mathcal{J}_i\}$, where $\mu_{jt}^{(i)}(x)$ takes value 1 if a request for item $j \in \mathcal{J}_i$ at time t when system inventory x is accepted, and zero otherwise. The set of admissible policies for Agent i is

  $$\mathcal{U}^{(i)}(x_t) = \{\mu_{jt}^{(i)}(x_t) \in \{0,1\}, j \in \mathcal{J}_i : A_j \mu_{jt}^{(i)}(x_t) \leq x_t\}$$

  meaning that requests can only be accepted if there is sufficient inventory.
- Agent i knows the probability q_{jt} that an item $j \in \mathcal{J}_i$ in his bundle will be requested at time t, and hence the probability $q_{jt}\mu_{jt}^{(i)}(x_t)$ that the associated resources A_j ($j \in \mathcal{J}_i$) are used;
- Agent i has a model of the probability that shared resources are consumed by another agent but it may be mis-specified. Specifically, if $\eta_{jt}^{(i)}(x_t)$ denotes the probability *assumed by Agent i* that resources A_j ($j \in \mathcal{J}_{-i}$) are consumed by

another agent, then $\eta_{jt}^{(i)}$ is typically not equal to the actual probability that this event occurs[1].

- Every Agent correctly assumes that at most one bundle from any Agent can be requested in any given period.

Under these assumptions, each agent has a model in which the accept/reject decisions for bundles in his portfolio is the control variable, the probability that he receives a request for an item in his/her bundle is accurately specified, but the probability that resources are consumed by other agents is generally mis-specified. Mathematically, Agent i adopts the demand model

$$x_{t+1} = x_t - \sum_{j \in \mathcal{J}_i} A_j \mu_{jt}^{(i)}(x) \widetilde{d}_{jt} - \sum_{j \in \mathcal{J}_{-i}} A_j \widetilde{d}'_{jt},\ t = 1, \cdots, T \tag{2.6}$$

where $\widetilde{d}_{jt},\ j \in \mathcal{J}_i$ is a binomial random variable which takes the value 1 if item j in Agent i's portfolio is requested but is otherwise 0, $\widetilde{d}'_{jt}$ is also binomial taking value 1 when inventory $A_j,\ j \in \mathcal{J}_{-i}$ is requested by another agent but is otherwise 0, and $\mu_{jt}^{(i)}(x_t)$ is Agent i's accept/reject policy for items in his inventory. For every bundle $j \in \mathcal{J}_i \cup \mathcal{J}_{-i}$, Agent i specifies probabilities

$$\begin{aligned} q_{jt} &= P[\widetilde{d}_{jt} = 1,\ \widetilde{d}_{kt} = 0\ \forall\, k \in \mathcal{J}_i/\{j\},\ \widetilde{d}'_{kt} = 0\ \forall\, k \in \mathcal{J}_{-i}],\ \text{if } j \in \mathcal{J}_i, \\ \eta_{jt}^{(i)}(x_t) &= P[\widetilde{d}'_{jt} = 1,\ \widetilde{d}_{kt} = 0\ \forall\, k \in \mathcal{J}_i],\ \text{if } j \in \mathcal{J}_{-i}. \end{aligned}$$

We assume that q_{jt} is correctly specified while $\eta_{jt}^{(i)}(x_t)$ may be mis-specified.

2.3.2 Transfer Contracts and Revenue Sharing

System inefficiencies caused by the selfish use of shared resources generally occur if the incentives of decentralized agents are not properly aligned with those of the system. To address this issue, we introduce the notion of a transfer contract which sets a price for using shared resources that is paid by the allocating agent and transferred to others as compensation.

Let $R_i(t, x, A_j)$ denote the payment received by Agent i whenever resources A_j ($j \in \mathcal{J}_{-i}$) are consumed by any other agent. It follows that if Agent i consumes A_j

[1]For example, if (a) item $j \in \mathcal{J}_{-i}$ happens to be controlled by Agent k, (b) the probability that Agent k receives a request is $q_{jt}^{(k)}$, and (c) Agent k has policy $\mu_{jt}^{(k)}(x_t)$, then the true probability that resources A_j are consumed is $q_{jt}^{(k)} \mu_{jt}^{(k)}(x_t)$, which we assume to be known by Agent k. However, there is typically no reason to assume that Agent i knows any of this information (or even the identity of the agent(s) who may use A_j), so we allow for the possibility that $\eta_{jt}^{(i)}(x_t) \neq q_{jt}^{(k)} \mu_{jt}^{(k)}(x_t)$ and that Agent i's model may be mis-specified.

$(j \in \mathcal{J}_i)$, that his net payout to other agents is

$$R_{-i}(t, x, A_j) \triangleq \sum_{i' \neq i} R_{i'}(t,\, x,\, A_j). \tag{2.7}$$

For Agent i, all that is relevant is the pair of functions

$$[R_{-i}(t,\, x,\, A_j),\, R_i(t,\, x,\, A_k),\; j \in \mathcal{J}_i,\; k \in \mathcal{J}_{-i}],$$

specifying the payment $R_{-i}(t,\, x,\, A_j)$ that needs to be made whenever he makes an allocation, and the payment $R_i(t,\, x,\, A_k)$ that is received whenever resources A_k are consumed by another agent. We say that the set of contracts

$$R(t,\, x) = \Big\{[R_{-1}(t,\, x,\, A_j),\, R_1(t,\, x,\, A_j)], \cdots,\, [R_{-|\mathcal{I}|}(t,\, x,\, A_j),\, R_{|\mathcal{I}|}(t,\, x,\, A_j)],\; j \in \mathcal{J}\Big\}$$

is *admissible* if it satisfies the condition (2.7). Transfer contracts change the expected profits and behavior of each agent. When they are admissible, payments are internal transfers between agents.

2.3.3 Decentralized Control

For a given set of revenue transfer contracts $[R_{-i}(t,\, x,\, A_j),\, R_i(t,\, x,\, A_j)]$, Agent i maximizes his expected profit conditional on his potentially mis-specified model for resource consumption

$$(C_i) \begin{cases} \max\limits_{\mu^{(i)}(\cdot)} \mathbb{E} \sum\limits_{t=1}^{T} \Big\{ \sum\limits_{j \in \mathcal{J}_i} \Big[r_j - R_{-i}(t, x_t, A_j) \Big] \mu_{jt}^{(i)}(x_t) \widetilde{d}_{jt} + \sum\limits_{j \in \mathcal{J}_{-i}} R_i(t, x_t, A_j) \widetilde{d}_{jt} \Big\} \\ \text{subject to:} \\ x_{t+1} = x_t - \sum\limits_{j \in \mathcal{J}_i} A_j \mu_{jt}^{(i)}(x_t) \widetilde{d}_{jt} - \sum\limits_{j \in \mathcal{J}_{-i}} A_j \widetilde{d}_{jt},\; x_0 = \mathcal{L}, \\ \mu_{jt}^{(i)}(x_t) \in \mathcal{U}^{(i)}(x). \end{cases} \tag{2.8}$$

The objective consists of two terms. The first $r_j - R_{-i}(t, x_t, A_j)$ is the net revenue received by Agent i when he accepts a request for one unit of bundle $j \in \mathcal{J}_i$ consisting of income r_j from the sale net the cash transfers $R_{-i}(t, x_t, A_j)$ to the other agents. The second $R_i(t, x_t, A_j)$ is the cash transfer received by Agent i whenever A_j $(j \in \mathcal{J}_{-i})$

is consumed by another agent. The dynamic programming equation for Agent i is

$$\begin{cases} V_i(t,x;R) = \max\limits_{\mu^{(i)}} \sum\limits_{j\in\mathcal{J}_i} q_{jt}\Big[\mu_{jt}^{(i)}(r_j - R_{-i}(t,x,A_j)) + V_i(t+1, x - A_j\mu_{jt}^{(i)})\Big] \\ \quad + \sum\limits_{j\in\mathcal{J}_{-i}} \eta_{jt}^{(i)}(x)\Big[R_i(t,x,A_j) + V_i(t+1, x - A_j)\Big] \\ \quad + \Big[1 - \sum\limits_{j\in\mathcal{J}_i} q_{jt} - \sum\limits_{j\in\mathcal{J}_{-i}} \eta_{jt}^{(i)}(x)\Big] V_i(t+1,x), \\ V_i(T+1,x;R) = 0. \end{cases}$$

Defining

$$\Delta V_i(t+1,x,A_j;R) \triangleq V_i(t+1,x;R) - V_i(t+1,x-A_j;R),$$

the dynamic programming equations for (2.8) can be written as

$$\begin{cases} V_i(t,x;R) = \max\limits_{\mu^{(i)}} \sum\limits_{j\in\mathcal{J}_i} q_{jt}\mu_{jt}^{(i)}\Big[r_j - R_{-i}(t,x,A_j) - \Delta V_i(t+1,x,A_j;R)\Big] \\ \quad + \sum\limits_{j\in\mathcal{J}_{-i}} \eta_{jt}^{(i)}(x)\Big[R_i(t,x,A_j) - \Delta V_i(t+1,x,A_j;R)\Big] + V_i(t+1,x;R), \\ V_i(T+1,x;R) = 0. \end{cases} \tag{2.9}$$

We emphasize that Agent i's model $\eta_{jt}^{(i)}(x)$ for the probability that resources A_j are consumed by another agent may be mis-specified.

2.3.4 Weak Duality

For every admissible contract R, each agent formulates a potentially mis-specified stochastic dynamic resource allocation problem (2.8) and solves for an optimal policy $\mathbf{u}_i^* = \{\mu_{jt}^{(i)*}(x), \forall j \in \mathcal{J}_i\}$. These policies are then implemented in the real world by each agent, where requests occur according to the true probabilities q_{jt} (2.1) that were used when formulating the centralized problem. The expected revenue for the aggregate system under the set of decentralized policies $\{\mathbf{u}_1^*, \cdots, \mathbf{u}_{|\mathcal{I}|}^*\}$ is given by

$$\begin{cases} \hat{V}(t,x;R) \triangleq \mathbb{E} \sum\limits_{s=t}^{T} \sum\limits_{i\in\mathcal{I}} \sum\limits_{j\in\mathcal{J}_i} r_j \mu_{js}^{(i)*}(x_s)\widetilde{d}_{js} \\ \text{subject to:} \\ x_{s+1} = x_s - \sum\limits_{i\in\mathcal{I}} \sum\limits_{j\in\mathcal{J}_i} A_j \mu_{js}^{(i)*}(x_s)\widetilde{d}_{jt}, \quad \forall\, t \le s \le T, \quad x_t = x, \end{cases} \tag{2.10}$$

where the distribution of $\widetilde{d}_{jt}$ is given by (2.1) and coincides with the probability adopted by the centralized decision maker. Whenever Agent i accepts a request, he is credited r_i and makes payment $R_{-i}(t,\, x,\, A_j)$ which, by (2.7), is divided amongst the other agents in the system with portion $R_{i'}(t,\, x,\, A_j)$ going to Agent i'. Observing that revenue sharing is nothing but an internal cash transfer, it follows that the net system profit equals the sum of profits by each agent. Clearly, it is of interest how the system profit $\hat{V}(t,\, x;\, R)$ under decentralized control relates to the optimal system profit given by the value function $V(t,\, x)$ of the centralized problem (2.3). The following result comes immediately from the observation that the decentralized policies $\{\mathbf{u}_1^*, \cdots, \mathbf{u}_{|\mathcal{I}|}^*\}$ are admissible though not generally optimal for the centralized problem.

Proposition 2.3.1 (Weak Duality). *Let R be an arbitrary transfer contract, $\{\mathbf{u}_1^*, \cdots, \mathbf{u}_{|\mathcal{I}|}^*\}$ the optimal decentralized policies under this contract, $\hat{V}(t, x; R)$ the resulting expected system revenue (2.10), and $V(t,\, x)$ the value function for the centralized agent (2.3). Then*

$$\hat{V}(t, x; R) \leq V(t, x), \quad \forall t, x. \tag{2.11}$$

Several questions are immediate:

- What choice of contracts "maximize" the efficiency of the decentralized system? Can contracts R be found under which decentralized agents achieve centralized efficiency?
- What is the impact of model mis-specification on efficiency of decentralized allocation?

2.4 Optimal Transfer Contracts

We begin by proposing conditions that optimal contracts *should* satisfy. Existence of contracts satisfying these conditions and the optimality of these contracts will then be established.

2.4.1 Optimality Conditions: Conjecture

Let

$$R(t,\, x) = \Big\{[R_{-1}(t,\, x,\, A_j),\, R_1(t,\, x,\, A_j)], \cdots, [R_{-|\mathcal{I}|}(t,\, x,\, A_j),\, R_{|\mathcal{I}|}(t,\, x,\, A_j)],\ j \in \mathcal{J}\Big\}$$

be an arbitrary admissible transfer contract and

$$V_1(t, x; R), \cdots, V_{|\mathcal{I}|}(t, x; R) \tag{2.12}$$

denote the value functions for each agent obtained by solving the decentralized problems (2.8) under R. $V_i(t,\, x;\, R)$ is Agent i's valuation of the shared inventory x at time t under contract R and $V_i(t+1,\, x;\, R) - V_i(t+1,\, x - A_j;\, R)$ is the opportunity cost to Agent i of losing inventory A_j. Observe that each agent's value function and opportunity cost depend on his mis-specified model for resource consumption.

A contract $R(t,\, x)$ defines cash transfers between agents whenever inventory is consumed; specifically, $R_{-i}(t,\, x,\, A_j)$ is the cost to Agent i of using resources A_j which is divided amongst other agents with the portion $R_{i'}(t,\, x,\, A_j)$ going to Agent i'. In this light, contracts define compensation payments and it is natural to expect that optimal contracts compensate each agent his opportunity cost for the inventory just consumed. This leads to the conjecture that optimal admissible contracts should satisfy

$$\left\{ \begin{array}{lcl} R_{-i}(t,\, x,\, A_j) & = & \sum_{i' \neq i} \Big[V_{i'}(t+1, x; R) - V_{i'}(t+1, x - A_j; R) \Big], \\ R_i(t,\, x,\, A_j) & = & V_i(t+1, x; R) - V_i(t+1, x - A_j; R). \end{array} \right. \tag{2.13}$$

This condition is a complicated system of implicit equations for the contract R. Several issues need to be resolved.

- Does a contract satisfying conditions (2.12)-(2.13) exist?
- What is the efficiency loss relative to centralized optimality when a contract satisfying (2.12)-(2.13) is implemented?
- What is the impact of model mis-specification? Does it minimize efficiency loss in some sense, even if the models adopted by each agent are mis-specified?
- Can a solution of (2.12)-(2.13) be computed without decentralized agents having to share information about their own models (e.g. request probabilities) in order to solve a centralized problem?

2.4.2 Existence

Let $V(t,\, x)$ denote the value function for the centralized problem (2.3), $V_i(t,\, x)$ the solution of the recursive equation

$$\left\{ \begin{array}{l} V_i(t,\, x) = \max_{\mu} \sum_{j \in \mathcal{J}_i} q_{jt} \mu_{jt}^{(i)} (r_j - \Delta V(t+1, x, A_j)) + V_i(t+1, x), \\ V_i(T+1,\, x) = 0. \end{array} \right. \tag{2.14}$$

and

$$\begin{cases} R_{-i}(t,x,A_j) &= \sum_{i' \neq i} \Big[V_{i'}(t+1,x;R) - V_{i'}(t+1,x-A_j;R) \Big] \\ R_i(t,x,A_j) &= V_i(t+1,x;R) - V_i(t+1,x-A_j;R). \end{cases} \tag{2.15}$$

Observe that $V(t,\, x)$ appears in the RHS of (2.14). Clearly, the contract (2.15) is admissible in that it satisfies the condition (2.7), and can be computed (at least in principle) by solving the centralized problem for $V(t,\, x)$ and the recursive equations (2.14) for $V_i(t,\, x)$. We now show that contract (2.14)-(2.15) satisfies the conjectured optimality conditions (2.12)-(2.13). The following preliminary result is required.

Proposition 2.4.1. *Let $V(t,\, x)$ denote the value function for the centralized problem (2.3) and $V_1(t,\, x), \cdots, V_{|\mathcal{I}|}(t,\, x)$ be defined in (2.14). Then $V(t,x) = \sum_{i\in\mathcal{I}} V_i(t,x)$.*

Proof. We shall prove this by induction. At $t = T+1$, the claim holds trivially. Suppose the claim holds for time period $t+1$, namely

$$V(t+1,x) = \sum_{i\in\mathcal{I}} V_i(t+1,x).$$

Since $V(t,\, x)$ solves the dynamic programming equation (2.4), it also solves

$$\begin{cases} V(t,\, x) = \max_{\mu^{(i)}} \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}_i} q_{jt}\mu_{jt}^{(i)}(r_j - \Delta V(t+1,x,A_j)) + V(t+1,x), \\ V(T+1,\, x) = 0, \end{cases} \tag{2.16}$$

where the RHS of this equation comes from decomposing the RHS of (2.4) over the bundles managed by different agents. On the other hand, summing (2.14) over i gives

$$\begin{aligned} \sum_{i\in\mathcal{I}} V_i(t,x) &= \sum_{i\in\mathcal{I}} \max_{\mu^{(i)}} \sum_{j\in\mathcal{J}_i} q_{jt}\mu_{jt}^{(i)}(r_j - \Delta V(t+1,x,A_j)) + \sum_{i\in\mathcal{I}} V_i(t+1,x) \\ &= \max_{\mu^{(i)}} \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}_i} q_{jt}\mu_{jt}^{(i)}(r_j - \Delta V(t+1,x,A_j)) + V(t+1,\, x), \end{aligned}$$

where the second equality follows from the induction hypothesis and the separability of the terms involving $\mu_{jt}^{(i)}$. Comparing with (2.16) it follows that

$$\sum_{i\in\mathcal{I}} V_i(t,x) = V(t,\, x),$$

and our result follows. □

We show that (2.15) is a solution of the system of equations (2.12)-(2.13) by showing that the solution $V_i(t,\, x)$ of (2.14) is the value function for the decentralized problem (2.8) under the contract (2.14)-(2.15). To see this, observe (by Proposition 2.4.1) that

$$\Delta V(t+1,x,A_j) = \sum_{i\in\mathcal{I}} \Delta V_i(t+1,x,A_j) = \sum_{i'\neq i} \Delta V_{i'}(t+1,x,A_j) + \Delta V_i(t+1,x,A_j).$$

This implies that (2.14) is equivalent to

$$V_i(t,x) = \max_{\mu} \sum_{j\in\mathcal{J}_i} q_{jt}\mu_{jt}^{(i)} \Big[r_j - \sum_{i'\neq i} \Delta V_{i'}(t,x,A_j) - \Delta V_i(t+1,x,A_j)\Big] + V_i(t+1,x).$$

When the contract R is given by (2.14)-(2.15), this equation can be written as

$$\begin{cases} V_i(t,x) = \max\limits_{\mu^{(i)}} \sum\limits_{j\in\mathcal{J}_i} q_{jt}\mu_{jt}^{(i)} \Big[r_j - R_{-i}(t,x,A_j) - \Delta V_i(t+1,\, x,\, A_j)\Big] \\ \quad + \sum\limits_{j\in\mathcal{J}_{-i}} \eta_{jt}^{(i)}(x)\Big[R_i(t,x,A_j) - \Delta V_i(t+1,\, x,\, A_j)\Big] + V_i(t+1,\, x), \\ V_i(T+1,\, x) = 0. \end{cases} \tag{2.17}$$

Observing that this is nothing but the dynamic programming equation for the decentralized problem (2.8) under contract (2.14)-(2.15), we can now say the following:

- The solution $V_i(t,\, x)$ of (2.14) equals the value function for the decentralized problem for Agent i under the contract (2.14)-(2.15). It follows that the contract (2.14)-(2.15) is a solution of the system of equations (2.12)-(2.13).
- The maximizer in the RHS of (2.14) is also the maximizer in the RHS of the dynamic programming equation (2.17) under contract (2.14)-(2.15). It follows that the maximizer in (2.14) defines the optimal allocation decision for Agent i under the contract (2.14)-(2.15).
- Under transfer contract (2.14)-(2.15), the second term in (2.17) is always zero. It follows that the value function of the decentralized problem (2.8) as well as the associated optimal allocation policy do not depend on Agent i's specification of the probability $\eta_{jt}^{(i)}(x)$ that another agent consumes inventory A_j.

We summarize these observations as follows.

Proposition 2.4.2. *Let the transfer contract R be defined by (2.14)-(2.15). Then R is a solution of the system of equations (2.12)-(2.13). Under this contract, the value function of Agent i's problem (2.8) is also the solution of (2.14) and the optimal allocation policy for Agent i is*

$$\mu_{jt}^{(i)*}(x) = \begin{cases} 1, & \textit{if } r_j \geq R_{-i}(t,x,A_j) + \Delta V_i(t+1,x,A_j), \\ 0, & \textit{otherwise.} \end{cases}$$

Both the value function $V_i(t, x)$ and the optimal policy $\mathbf{u}_i^* = \{\mu_{jt}^{(i)*}(x), j \in \mathcal{J}_i\}$ are independent of Agent i's specification in his model (2.8) of the probability $\eta_{jt}^{(i)}(x)$ that inventory A_j is consumed by another agent.

2.4.3 Verification of Optimality

Proposition 2.4.2 tells us that the contract defined in (2.14)-(2.15) satisfies the conjectured optimality conditions (2.12)-(2.13), and that the value functions of the decentralized agents under this contract are insensitive to mis-specification of the probability of consumption by other agents. We now show that the contracts (2.14)-(2.15) are optimal in that they achieve equality in (2.11) and that the decentralized policies are optimal for the integrated system.

Theorem 2.4.1 (Strong duality)**.** *Let $(R, (V_1(t, x), \cdots, V_{|\mathcal{I}|}(t, x)))$ denote the solution of the system (2.14)-(2.15) and $\mathbf{u}_i^*$ be the optimal allocation policy for Agent i's problem (2.8) under this contract. Then*

1. *R solves the system of equations (2.12)-(2.13) and $V_i(t, x)$ equals the value function for Agent i's problem (2.8) under this contract.*

2. *The collection of decentralized policies $\{\mathbf{u}_1^*, ..., \mathbf{u}_{|\mathcal{I}|}^*\}$ under contract (2.14)-(2.15) is optimal for the centralized problem (2.3) and the system profit under these decentralized policies equals the optimal profit for the centralized agent;*

3. *The value function and optimal allocation policy $\mathbf{u}_i^*$ for Agent i are independent of his specification $\eta_{jt}^{(i)}(x)$ of the probability that resources A_j $(j \in \mathcal{J}_{-i})$ are consumed by other agents in (2.8).*

Proof. Properties (1) and (3) were shown in Proposition 2.4.2.

Suppose that the transfer contract is given by (2.14)-(2.15). Then the optimal allocation policy, by Proposition 2.4.2, is given by

$$\begin{aligned}\mathbf{u}_i^* &= \arg\max_{\mu^{(i)}} \sum_{j\in\mathcal{J}_i} q_{jt}\mu_{jt}^{(i)}\Big[r_j - R_{-i}(t, x, A_j) - \Delta V_i(t+1, x, A_j)\Big] \\ &= \arg\max_{\mu^{(i)}} \sum_{j\in\mathcal{J}_i} q_{jt}\mu_{jt}^{(i)}\Big[r_j - \Delta V(t+1, x, A_j)\Big],\end{aligned}$$

where $V(t, x)$ is the value function for the centralized agent which solves the dynamic programming equation (2.4). This implies that the collection of decentralized policies under contract (2.14)-(2.15) satisfies

$$\{\mathbf{u}_1^*, ..., \mathbf{u}_{|\mathcal{I}|}^*\} = \arg\max_{\mu^{(i)}} \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}_i} q_{jt}\mu_{jt}^{(i)}(r_j - \Delta V(t+1, x, A_j)).$$

Property (2) now follows from the observation that this is precisely the optimal policy for the centralized problem as defined by (2.4). □

Though transfer contracts play a role analogous to shadow prices in classical resource allocation problems, it is notable that strong duality holds without any assumptions of convexity.

2.5 Decentralized Computation of Optimal Contracts

Theorem 2.4.1 characterizes transfer contracts under which decentralized agents optimally choose the centrally optimal policies. We now turn to the question of computing these contracts. One approach is to solve the system of equations (2.14)-(2.15) directly, but this can only be done if the demand probabilities q_{jt} of all bundles and the value function of the centralized problem $V(t,\, x)$ are known. This is disquieting because our motivation for studying the decentralized problem was the argument that agents with this information typically do not exist.

2.5.1 Iterative Decentralized Algorithm

The following algorithm is motivated by the above considerations and can be viewed as an iterative approach for solving the implicit system of equations (2.12)-(2.13). (Recall from Theorem 2.4.1 that the optimal contract is a solution of this system). In each iteration, decentralized agents solve their mis-specified optimization problems (2.8) conditional on some suboptimal transfer contract. These contracts are then updated locally by each of the agents and exchanged, and the process repeats. While it is natural to ask whether the algorithm converges and whether the limiting contract is optimal, which we address in Section 2.6, it is important to recognize that the algorithm can be implemented without an all knowing centralized agent and does not require agents to exchange private information about demand probabilities or to even know how many other agents there are in the system. All that is exchanged are updated transfer contracts and the algorithm allows for the possibility that the decentralized problems may be mis-specified.

Algorithm 2.5.1. *(Iterative Decentralized Algorithm)*

Initialize: Set $k = 1$, *and* $R_i^1(t,x,A_j) = 0, R_{-i}^1(t,x,A_j) = 0$.

Step 1: Given k, *and* $R_i^k(t,x,A_j), R_{-i}^k(t,x,A_j)$

- *Each agent solves his own decentralized problem (2.8) and computes his value function* $V_i^k(t,x) = V_i(t,\, x;\, R^k)$ *by solving (2.9);*

Airline	Leg (capacity)	Itinerary (fare)
1	A (10)	A ($250), AB ($400)
2	B (10)	B ($250), BC ($400)
3	C (10)	C ($250)

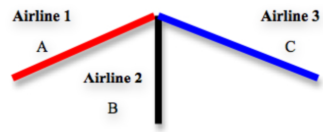

Figure 2.1: A three-airline alliance with three legs and five itineraries: Airline 1 and Airline 2 operates both intraline (A and B respectively) and interline itineary (AB and BC respectively), while Airline 3 only operates intraline itinerary (C).

- *Stop if, a satisfactory level of precision has been reached,*

$$\sup_{i,t,x} |V_i^k(t,x) - V_i^{k-1}(t,x)| \leq \epsilon;$$

otherwise, each agent updates his transfer contract,

$$R_i^{k+1}(t,x,A_j) = V_i^k(t+1,x) - V_i^k(t+1,x-A_j).$$

- *Each agent communicates the updated transfer contract* $R_i^{k+1}(t,x,A_j)$ *to the system.*

- *System synthesize* $R_{-i}^{k+1}(t,x,A_j) = \sum_{i' \neq i} R_i^{k+1}(t,x,A_j)$*, and broadcast them back to all agents.*

Step 2: Set k *to* $k+1$ *and return to Step 1.*

2.5.2 Numerical Example

We now illustrate our decentralized resource allocation framework with an example from airline alliance revenue management. Figure 2.1 shows an airline alliance consisting of three agents (Airline 1/2/3). The alliance has three resources (flight-leg A/B/C), and markets five bundles (itinerary A/B/C/AB/BC). The arrival probability for itineraries are set such that the network has an overall load factor of 1.33, where the load factor,

$$\alpha = \frac{\sum_{ljt} q_{jt} a_{lj}}{\sum_l x_{l1}},$$

is the ratio of expected resource consumption to the initial inventory level. Lastly the planning horizon has $T = 30$ periods.

Convergence

Figure 2.2 shows the convergence of Algorithm 2.5.1 (both in terms of error 1 that is evaluated under sup-norm over the entire state space, as well as error 2 that is

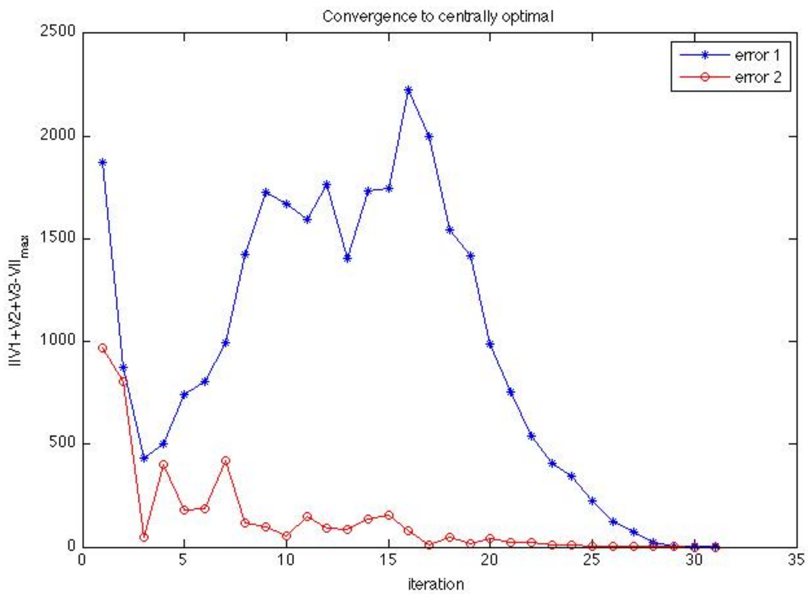

Figure 2.2: Convergence of Algorithm 2.5.1 with two types of errors: error 1 = $\sup_{t,x} || \sum_i V_i^k(t,x) - V(t,x)||$, error 2 = $| \sum_i V_i^k(1,10) - V(1,10)|$.

evaluated only at the initial time and maximum capacity) where the joint value function converges to the centrally optimal solution (we solve the benchmark centralized optimal value function separately). If we look at error 1, which is the worst case error, the convergence is not very smooth. However, if we look at error 2, which is the error associated with the start of the planning horizon, it converges much faster and smoother. That is an nice property for practical implementation, as the initial decisions only depend on the value functions at the initial time, which means we could run the algorithm dynamically over time and retain only the value functions at the initial time, and only a small number of iterations is required to achieve good accuracy. Figure 2.3 shows the decomposition of the the centrally optimal value function into the value functions of the three individual airlines, and we see that the value functions decrease monotonically over time as airline inventory is perishable.

Transfer contracts and the impact of network topology

Note that our example has a special network structure, such that some airlines are directly connected (e.g. Airline 1 and 2), while some are only indirectly connected (e.g. Airline 1 and 3 via 2). We would like to examine the dependence of the revenue transfers on the underlying network topology. We define a first-order transfer as a payment to a directly connected partner (e.g. Airline 1 to 2), and a second-order transfer as a payment to an indirectly connected partner (e.g. Airline 1 to 3).

Table 2.1 shows the break-down of the net revenue transfer R into first and second

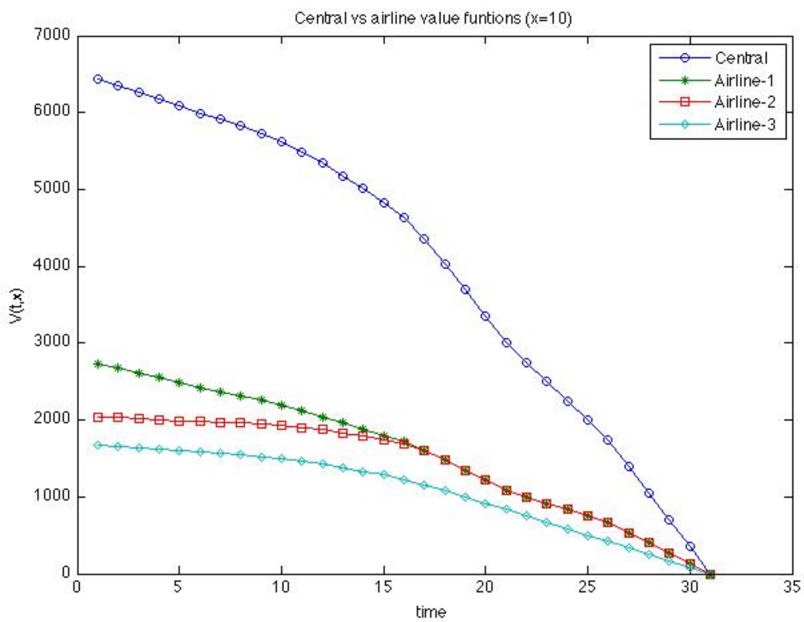

Figure 2.3: Central optimal value function $V(t,x)$ vs individual airlines optimal value function $V_1(t,x), V_2(t,x), V_3(t,x)$: decays monotonically over time as airline inventory is perishable.

order transfers from a simulated sample path of the demand and revenue realization. For example, when Airline 3 makes the first sale, $32 is transferred to Airline 2 (first order), while -$15 is transferred to Airline 1 (second order). The most surprising feature is that revenue transfers can take on negative values, meaning that having less inventory can be sometimes beneficial for some airline. For example, Airline 1 is always willing to subsidize the sale by taking a negative revenue transfer whenever Airline 3 makes a sale (which is second order since Airline 1 and -3 are only indirectly connected). The reason is that when Airline 3 makes a sale of itinerary C, Airline 2 will have less opportunity to sell itinerary BC, thus increasing the opportunity for Airline 1 to sell itinerary AB. Clearly, network topology can have a strong impact on the distribution of revenue transfer, and one should exploit such structural properties in designing practical transfer schemes.

Selling Airline	R	1st order R	2nd order R
3	17	32	-15
1	-13	-18	5
2	22	68	-46
3	29	49	-20
3	41	64	-23
3	45	64	-19
2	237	54	183
1	93	120	-27
1	94	129	-35
1	111	151	-40
3	0	16	-16
1	-21	-28	7
3	11	30	-19
1	-29	-33	4

Table 2.1: Decomposed revenue transfers from a simulated sample path.

2.6 Convergence

Transfer contracts R specify prices that are paid by agents when consuming shared resources and as such are related to Lagrange multipliers/shadow prices in classical resource allocation problems. In this light, Algorithm 2.5.1 is analogous to a dual type approach for updating prices. We now present results that guarantee convergence of the Algorithm 2.5.1 to the optimal contract. As in Theorem 2.4.1 convexity is not required to guarantee convergence, and convergence is robust to agent-level mis-specification of the demand probabilities of other agents.

2.6.1 Statement of Main Results

Theorem 2.6.1. *Let $V_i(t,x)$ be the value function for Agent i's decentralized problem (2.9) under the optimal contract (2.15), and the sequence of $\{V_i^k(t,x)\}$ be computed by Algorithm 2.5.1. Then $\{V_i^k(t,x)\}$ converges strongly to $V_i(t,x)$, namely*

$$\lim_{k\to\infty} \|V_i^k(t,x) - V_i(t,x)\| = 0, \quad \forall i \in \mathcal{I}.$$

Suppose $\mathbf{u}_i^k = \{\mu_{jt}^k(x), j \in \mathcal{J}_i\}$ [2] is the optimal policy for Agent i obtained in the k^{th} iteration, and let

$$\begin{cases} \hat{V}^k(t,x) \triangleq \mathbb{E}\sum_{s=t}^{T}\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_i} r_j\mu_{js}^k(x_s)\widetilde{d}_{js} \\ \text{subject to:} \\ x_{s+1} = x_s - \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_i} A_j\mu_{js}^k(x_s)\widetilde{d}_{jt}, \quad \forall\, t \le s \le T, \quad x_t = x, \end{cases} \tag{2.18}$$

denote the system profit under the set of policies $\{\mathbf{u}_1^k, ..., \mathbf{u}_{|\mathcal{I}|}^k\}$ obtained from the k^{th} iteration of Algorithm 2.5.1. The following result guarantees convergence of $\hat{V}^k(t,x)$ to the value function $V(t,x)$ of the centralized agent.

Theorem 2.6.2. *There exist constants $C, M > 0$ such that*

$$\|\hat{V}^k(t,x) - V(t,x)\| \le \frac{C(MT)^{k+1}}{M(k+1)!}, \ \forall k \ge 1.$$

It follows that

$$\lim_{k\to\infty} \|\hat{V}^k(t,x) - V(t,x)\| = 0.$$

We now turn to a proof of Theorems 2.6.1 and 2.6.2. To ease the notation, we assume in the proof that $\{\eta_{jt}^{(i)}(x) \equiv 0, \forall j \notin \mathcal{J}_i\}$ in Agent i's model; i.e. that Agent i (incorrectly) assumes in his/her model that inventory can not be depleted by other agents. We emphasize that this simplification is being made to reduce the length of equations/plethora of subscripts and to improve the clarity of the proof, and that all results holds for the more general case when Agent i adopts non-zero values for the demand probabilities of other agents. We note in particular that each agent's model is still mis-specified under this simplification, so we are not assuming away an essential feature of our system/result in making this assumption. Under these assumptions, the decentralized model of Agent i

[2] To ease the notation, we use $\mu_{jt}^k(x)$ instead of $\mu_{jt}^{(i)k}(x)$ in the remaining analysis, with the index $j \in \mathcal{J}_i$ indicating that the control policy belongs to Agent i.

(2.8) simplifies to

$$(\hat{C}_i) \begin{cases} \max_{\mu(\cdot)} \mathbb{E} \sum_{t=1}^{T} \sum_{j \in \mathcal{J}_i} \Big[r_j - R_{-i}(t, x, A_j) \Big] \mu_{jt}(x_t) \widetilde{d}_{jt} \\ \text{subject to:} \\ x_{t+1} = x_t - \sum_{j \in \mathcal{J}_i} A_j \mu_{jt}(x) \widetilde{d}_{jt}, \; x_0 = \mathcal{L}, \\ \mu_{jt}(x_t) \in \mathcal{U}(x_t). \end{cases} \tag{2.19}$$

For which the dynamic programming equation is

$$\begin{cases} V_i(t,\, x) = \max_{\mu} \sum_{j \in \mathcal{J}_i} q_{jt} \mu_{jt} \Big[r_j - R_{-i}(t, x, A_j) - \Delta V_i(t+1,\, x,\, A_j) \Big] + V_i(t+1,\, x), \\ V_i(T+1, x) = 0. \end{cases} \tag{2.20}$$

2.6.2 Preliminaries

The well known inequality

$$|\max_x f(x) - \max_x g(x)| \le \max_x |f(x) - g(x)| \tag{2.21}$$

is used repeatedly in the proof. The following technical result is also required.

Lemma 2.6.1.

$$\sum_{s=t}^{T} \frac{(T-s)^{k-1}}{(k-1)!} \le \frac{(T-t+1)^k}{k!}. \tag{2.22}$$

Proof. This result follows from the observation that

$$\begin{aligned} \sum_{s=t}^{T} \frac{(T-s)^{k-1}}{(k-1)!} &\le \sum_{s=t}^{T} \int_s^{s+1} \frac{(T-u+1)^{k-1}}{(k-1)!} du, \\ &= \sum_{s=t}^{T} \Big[\frac{(T-s+1)^k}{k!} - \frac{(T-s)^k}{k!} \Big], \\ &= \frac{(T-t+1)^k}{k!}, \end{aligned}$$

where the first inequality basically says that the area under the step function in Figure 2.4 is bounded above by the area under the curve. □

Proposition 2.6.1. *There exists constants $B, M > 0$ such that for any agent $i \in \mathcal{I}$, we have*

$$\max_x |V_i^k(t,x) - V_i^{k-1}(t,x)| \le \frac{B[M(T-t+1)]^k}{Mk!}, \; \forall k \ge 1.$$

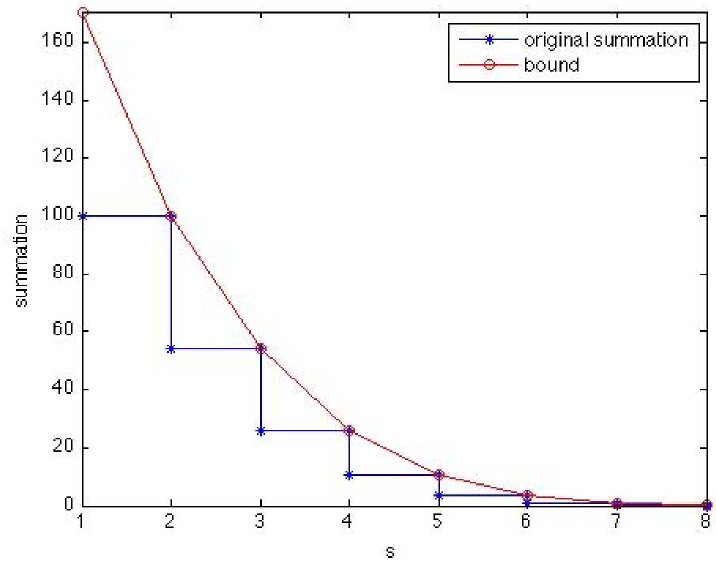

Figure 2.4: An upper bound for the summation (k = 5, t =1, T = 8)

It follows that

$$\|V_i^k - V_i^{k-1}\| = \sup_{t\in[1,T]} \max_x |V_i^k(t,x) - V_i^{k-1}(t,x)| \le \frac{B(MT)^k}{Mk!},$$

hence

$$\lim_{k\to\infty} \|V_i^k(t,x) - V_i^{k-1}(t,x)\| = 0.$$

Proof. Define,

$$W_i^k(t) \triangleq \max_x |V_i^k(t,x) - V_i^{k-1}(t,x)|, \quad \forall i \in \mathcal{I} \tag{2.23}$$

Assuming without loss of generality that $V_i^0(t,x) = 0$, we have

$$\begin{cases} V_i^1(t,x) \triangleq \max_{\mu(\cdot)} \mathbb{E} \sum_{s=t}^{T} \sum_{j\in\mathcal{J}_i} r_j \mu_{js}(x_s) \widetilde{d}_{js} \\ \text{subject to:} \\ x_{s+1} = x_s - \sum_{j\in\mathcal{J}_i} A_j \mu_{js}(x_s) \widetilde{d}_{jt}, \quad \forall\, t \le s \le T, \\ x_t = x, \end{cases} \tag{2.24}$$

since $R_{-i}^1(t,x,A_j) = \sum_{i'\neq i} \Delta V_{i'}^0(t,x,A_j) = 0$. Let $B = \max_{j\in\mathcal{J}}\{r_j\}$. Then B is an upper bound to the maximum revenue achievable in any time period by any agent solving (2.24) and it follows that

$$W_i^1(t) = \max_x |V_i^1(t,x)| \leq B(T-t+1), \quad \forall i \in \mathcal{I}. \tag{2.25}$$

Suppose now there is a k such that

$$W_i^{k-1}(t) = \max_x |V_i^{k-1}(t,x) - V_i^{k-2}(t,x)| \leq \frac{B[M(T-t+1)]^{k-1}}{M(k-1)!}, \tag{2.26}$$

where $M = 2(|\mathcal{I}|-1)$ is a constant. (By (2.25) this condition holds when $k=2$). Then

$$\begin{aligned}
&|V_i^k(t,x) - V_i^{k-1}(t,x)| \\
= \; & \Big| \max_{\mu(\cdot)} \Big\{ \mathbb{E} \sum_{s=t}^{T} \sum_{j \in \mathcal{J}_i} \Big[r_j - R_{-i}^k(t,x,A_j) \Big] \mu_{jt}(x_t) \widetilde{d}_{jt} \Big\} \\
& - \max_{\mu(\cdot)} \Big\{ \mathbb{E} \sum_{s=t}^{T} \sum_{j \in \mathcal{J}_i} \Big[r_j - R_{-i}^{k-1}(t,x,A_j) \Big] \mu_{jt}(x_t) \widetilde{d}_{jt} \Big\} \Big|, \\
\leq \; & \max_{\mu(\cdot)} \Big| \Big\{ \mathbb{E} \sum_{s=t}^{T} \sum_{j \in \mathcal{J}_i} \Big[r_j - R_{-i}^k(t,x,A_j) \Big] \mu_{jt}(x_t) \widetilde{d}_{jt} \Big\} \\
& - \Big\{ \mathbb{E} \sum_{s=t}^{T} \sum_{j \in \mathcal{J}_i} \Big[r_j - R_{-i}^{k-1}(t,x,A_j) \Big] \mu_{jt}(x_t) \widetilde{d}_{jt} \Big\} \Big|, \\
= \; & \max_{\mu(\cdot)} \Big| \mathbb{E} \sum_{s=t}^{T} \sum_{j \in \mathcal{J}_i} \sum_{i' \neq i} (-\Delta V_{i'}^{k-1}(s+1,x_s,A_j) + \Delta V_{i'}^{k-2}(s+1,x_s,A_j)) \mu_{js}(x_s) \widetilde{d}_{js} \Big|, \\
= \; & \max_{\mu(\cdot)} \Big| \mathbb{E} \sum_{s=t}^{T} \sum_{j \in \mathcal{J}_i} \sum_{i' \neq i} [(V_{i'}^{k-1}(s+1,x_s-A_j) - V_{i'}^{k-2}(s+1,x_s-A_j)) \\
& - (V_{i'}^{k-1}(s+1,x_s) - V_{i'}^{k-2}(s+1,x_s))] \mu_{js}(x_s) \widetilde{d}_{js} \Big|,
\end{aligned}$$

where the inequality follows from (2.21). It now follows that

$$\begin{aligned}
W_i^k(t) \;&\leq\; \sum_{s=t}^{T} \sum_{j \in \mathcal{J}_i} q_{jt} \sum_{i' \neq i} 2W_{i'}^{k-1}(s+1), \\
&\leq\; \sum_{s=t}^{T} M \frac{B[M(T-s)]^{k-1}}{M(k-1)!}, \\
&\leq\; \frac{BM^k}{M} \frac{(T-t+1)^k}{k!}, \\
&=\; \frac{B[M(T-t+1)]^k}{Mk!},
\end{aligned}$$

where the first inequality follows from the definition of $W_i^{k-1}(t)$ (see (2.26)), the second inequality follows from the definition of $M = 2(|\mathcal{I}|-1)$ together with the induction hypothesis (2.26) and the third inequality follows from (2.22) in Lemma 2.6.1. Taking supremums over $t \in [1, T]$ on both sides gives

$$\|V_i^k - V_i^{k-1}\| = \sup_{t\in[1,T]} W_i^k(t) \leq \frac{B(MT)^k}{Mk!},$$

and it follows that

$$\lim_{k\to\infty} \|V_i^k - V_i^{k-1}\| = 0, \quad \forall i \in \mathcal{I},$$

as claimed. □

2.6.3 Proof of Main Results

Proof of Theorem 2.6.1

For any $k \in \mathbb{Z}^+$ and $m \in \mathbb{Z}^+$,

$$\begin{aligned}
\|V_i^{k+m} - V_i^k\| \leq & \sum_{l=1}^{m} \|V_i^{k+l} - V_i^{k+l-1}\|, \\
\leq & \sum_{l=1}^{m} \frac{B(MT)^{k+l}}{M(k+l)!} \text{ (Proposition 2.6.1)}, \\
\leq & \frac{B(MT)^{k+1}}{M(k+1)!} \sum_{l=0}^{\infty} \frac{(MT)^l}{l!}, \\
= & \frac{B(MT)^{k+1}}{M(k+1)!} e^{-MT}, \quad \forall i \in I
\end{aligned}$$

which goes to zero as $k \to +\infty$. Therefore $\{V_i^k,\ k \geq 1\}$ is a Cauchy sequence under the sup-norm, so there exists a function $\hat{V}_i$ such that $\|V_i^k - \hat{V}_i\| \to 0$ as $k \to \infty$. On the other hand, we also know that V_i^k is the solution of the dynamic programming equation

$$\begin{cases} V_i^k(t,x) = \max\limits_{\mu} \sum\limits_{j\in\mathcal{J}_i} q_{jt}\mu_{jt}\Big[r_j - \sum\limits_{i'\neq i} \Delta V_{i'}^{k-1}(t+1,x,A_j) - \Delta V_i^k(t+1,x,A_j)\Big] + V_i^k(t+1,x), \\ V_i(T+1,\, x) = 0. \end{cases}$$

Taking limits on both sides gives

$$\begin{cases} \hat{V}_i(t,x) = \max\limits_{\mu} \sum\limits_{j\in\mathcal{J}_i} q_{jt}\mu_{jt}\Big[r_j - \sum\limits_{i'\neq i} \Delta \hat{V}_{i'}(t+1,x,A_j) - \Delta \hat{V}_i(t+1,x,A_j)\Big] + \hat{V}_i^k(t+1,x), \\ \hat{V}_i(T+1,\, x) = 0, \end{cases}$$

which is exactly the same dynamic programming equation as (2.20) under the optimal sharing contract (2.15). It follows that the limit $\hat{V}_i(t,\, x)$ equals the value function for Agent

i under the optimal transfer contract

$$\hat{V}_i(t,x) = V_i(t,x),$$

and that the value functions of the decentralized agents generated at each step of Algorithm 2.5.1 converges

$$\lim_{k\to\infty} \|V_i^k(t,x) - V_i(t,x)\| = 0, \quad \forall i \in \mathcal{I}.$$

Proof of Theorem 2.6.2

Since $\mathbf{u}_i^k = \{\mu_{jt}^k(x), j \in \mathcal{J}_i\}$ is the optimal policy for Agent i obtained at the k^{th} iteration, we have

$$\begin{cases} V_i^k(t,x) = \sum_{j\in\mathcal{J}_i} q_{jt}\mu_{jt}^k(x)\Big[r_j - R_{-i}^k(t,x,A_j) - \Delta V_i^k(t+1,x,A_j)\Big] + V_i^k(t+1,x), \\ V_i^k(T+1,x) = 0. \end{cases}$$

Defining $V^k(t,x) = \sum_{i\in\mathcal{I}} V_i^k(t,x)$, we have

$$\begin{aligned} V^k(t,x) &= \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_i} q_{jt}\mu_{jt}^k(x)(r_j - R_{-i}^k(t,x,A_j) - \Delta V_i^k(t+1,x,A_j)) + \sum_{i\in\mathcal{I}} V_i^k(t+1,x), \\ &= \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_i} q_{jt}\mu_{jt}^k(x)(r_j - R_{-i}^k(t,x,A_j) + \sum_{i'\neq i} \Delta V_{i'}^k(t+1,x,A_j) \\ &- \sum_{i\in\mathcal{I}} \Delta V_i^k(t+1,x,A_j)) + V^k(t+1,x). \end{aligned}$$

Since $\sum_{i'\neq i} \Delta V_{i'}^k(t+1,x,A_j) = R_{-i}^{k+1}(t,x,A_j)$, the above simplifies to

$$\begin{cases} V^k(t,x) = \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_i} q_{jt}\mu_{jt}^k(x)\Big[r_j - R_{-i}^k(t,x,A_j) + R_{-i}^{k+1}(t,x,A_j) \\ \quad -\Delta V^k(t+1,x,A_j)\Big] + V^k(t+1,x), \\ V^k(T+1,x) = 0. \end{cases} \tag{2.27}$$

It is clear that (2.27) is the dynamic programming equation of the stochastic control problem

$$\begin{cases} V^k(t,x) \triangleq \mathbb{E}\sum_{s=t}^{T}\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_i} (r_j - R_{-i}^k(t,x,A_j) + R_{-i}^{k+1}(t,x,A_j))\mu_{js}(x_s)^k \widetilde{d}_{js} \\ \text{subject to:} \\ x_{s+1} = x_s - \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_i} A_j\mu_{js}(x_s)^k(x_s)\widetilde{d}_{jt}, \quad \forall\, t \leq s \leq T, \\ x_t = x. \end{cases} \tag{2.28}$$

From (2.18) and (2.28) we have

$$\begin{aligned}
\left|\hat{V}^k(t,x) - V^k(t,x)\right| &= \left|\mathbb{E}\sum_{s=t}^{T}\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_i}\Big(-R^k_{-i}(t,x,A_j) + R^{k+1}_{-i}(t,x,A_j)\Big)\mu_{js}(x_s)^k\widetilde{d}_{js}\right|, \\
&= \Bigg|\mathbb{E}\sum_{s=t}^{T}\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_i}\sum_{i'\neq i}\Big[-\Big(V^{k-1}_{i'}(s+1,x_s) - V^{k-1}_{i'}(s+1,x_s-A_j)\Big) \\
&\quad +\Big(V^{k}_{i'}(s+1,x_s) - V^{k}_{i'}(s+1,x_s-A_j)\Big)\Big]\mu_{js}(x_s)\widetilde{d}_{js}\Bigg|, \\
&= \Bigg|\mathbb{E}\sum_{s=t}^{T}\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_i}\sum_{i'\neq i}\Big[-\Big(V^{k-1}_{i'}(s+1,x_s) - V^{k}_{i'}(s+1,x_s)\Big) \\
&\quad +\Big(V^{k-1}_{i'}(s+1,x_s-A_j) - V^{k}_{i'}(s+1,x_s-A_j)\Big)\Big]\mu_{js}(x_s)\widetilde{d}_{js}\Bigg|, \\
&\le \sum_{s=t}^{T}\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}_i}q_{jt}\sum_{i'\neq i}\Big[\Big|V^{k-1}_{i'}(s+1,x_s) - V^{k}_{i'}(s+1,x_s)\Big| \\
&\quad +\Big|V^{k-1}_{i'}(s+1,x_s-A_j) - V^{k}_{i'}(s+1,x_s-A_j)\Big|\Big], \\
&\le \sum_{s=t}^{T} M\frac{B[M(T-s)]^k}{Mk!}, \\
&\le \frac{B[M(T-t+1)]^{(k+1)}}{M(k+1)!},
\end{aligned}$$

hence,

$$\|\hat{V}^k(t,x) - V^k(t,x)\| \le \frac{B(MT)^{k+1}}{M(k+1)!}.$$

Also we have,

$$\begin{aligned}
\|V^k(t,x) - V(t,x)\| &\le \sum_{i\in\mathcal{I}}\|V^k_i(t,x) - V_i(t,x)\|, \\
&\le \sum_{i\in\mathcal{I}}\frac{B(MT)^{k+1}}{M(k+1)!}e^{-MT}, \\
&= (\frac{1}{2}M+1)\frac{B(MT)^{k+1}}{M(k+1)!}e^{-MT}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\|\hat{V}^k(t,x) - V(t,x)\| &\le \|\hat{V}^k(t,x) - V^k(t,x)\| + \|V^k(t,x) - V(t,x)\|, \\
&\le \frac{B(MT)^{k+1}}{M(k+1)!} + (\frac{1}{2}M+1)\frac{B(MT)^{k+1}}{M(k+1)!}e^{-MT}, \\
&= \frac{C(MT)^{k+1}}{M(k+1)!},
\end{aligned}$$

where $C = B[1 + (\frac{1}{2}M + 1)e^{-MT}]$ is a constant. Therefore, we have

$$\lim_{k\to\infty} \|\hat{V}^k(t,x) - V(t,x)\| = 0.$$

2.7 Summary

In this chapter we study a general class of stochastic dynamic resource allocation problems that has many real-world applications. We characterize transfer contracts that coordinate the system and show that centralized optimality is achieved even when the models adopted by decentralized agents are mis-specified. An iterative algorithm for decentralized computation of the optimal transfer contract is also proposed and shown to converge to the optimal transfer contract. Transfer contracts are analogous to shadow prices and the iterative algorithm to a dual update algorithm, so it is interesting that existence of transfer contracts that achieve centralized optimality (strong duality) and convergence to the optimal of the iterative algorithm can be guaranteed without assumptions of convexity. This is actually a special case of a substantially more general result, that will be explored elsewhere, that comes from our ability to characterize transfer contracts using dynamic programming arguments. We further illustrate our proposed framework with a case study in airline alliance revenue management. We show that the iterative algorithm converges strongly to the set of optimal transfer contracts, under which individual airlines' local pricing policies would maximize system revenue. We demonstrate that the amounts of revenue transfer can be highly nonlinear, change dynamically over time, and depend strongly on remaining network inventory and demand realization. Moreover, we observe that network topology can have a strong impact on the transfer contracts, such as resulting in negative amount of revenue transfer which is highly counter-intuitive. We believe that these insights could be exploited in constructing the optimal transfer contracts, such as to define a good initial set of transfer contracts and speed up convergence of the iterative algorithm.

Chapter 3

Approximation Methods for Decentralized Resource Allocation

3.1 Introduction

In the previous chapter we focus on characterizing the optimal contract that would lead to efficient coordination, such that locally optimal decisions made by decentralized agents coincide with the centrally optimal decisions that would be made if there exists an all-knowing central agent. We also show that the optimal contract can be constructed iteratively in a decentralized manner such that no private information needs to be revealed by local agents. Furthermore, optimality and convergence can be guaranteed even in the presence of agent-level model mis-specification, and no convexity is required. These are the central theoretical issues concerning decentralized control of stochastic dynamic multi-agent systems. However, on the practical side, an equally important issue is computational efficiency of such coordination mechanism. A practical-minded reader would notice in Algorithm 2.5.1, at each iteration each agent has to solve a stochastic dynamic control problem (2.9), which can quickly become intractable as the size of the problem increases. In this chapter, we seek to develop approximation methods and algorithms that can solve large-scale decentralized resource allocation problems fast and efficiently.

The computational issue is not just relevant for decentralized control, it is important also for centralized control and have previously received a lot of attentions in the literature. Despite the simple structure of the optimal policy in (2.5), it can be notoriously difficult to evaluate the value function $V(t,x)$ exactly, due to the well-known curse of dimensionality. Instead of trying to evaluate $V(t,x)$ directly, a simple but powerful framework has been proposed that bear the notion of bid-price control (Talluri and van Ryzin [36]). The idea is to approximate the value function by a (first-order) gradient approximation, and associate bid-price to each resource that captures its future opportunity cost. In the same spirit of the optimal policy (2.5), a customer request for a particular bundle is only accepted if there is enough inventory and the revenue received is greater than the combined bid-price of the resource required:

$$\begin{aligned} r_j &\geq V(t+1,x) - V(t+1,x-A_j), \\ &\approx (\Delta V(t+1,x))^T A_j, \\ &\approx \sum_{l=1}^{K} \pi_{l,t+1} a_{lj}, \end{aligned} \tag{3.1}$$

where $\pi_{l,t+1}$ denotes the bid-price for resource l at time $t+1$, and a_{lj} denotes the l's component of the resource vector A_j. Under the bid-price framework, a number of approximation methods and algorithms have been proposed as how to efficiently find bid-price $\pi_{l,t+1}$ that can approximate the value function well and thereafter be used to construct good control policy.

Most of the approximation methods and algorithms have been studied under the centralized setting, where having full information of the demand forecast and inventory status allows a central agent to accurately evaluate the bid-price for every resource. In particular, early work on computing bid-price is largely based on static linear programming approximations (Simpson [35], Williamson [41]). The static linear programming approach formulates a very simple one-stage planning problem, denoted as the primal problem. In the primal problem we decide how much resource to allocate to each and every bundle, subject to its expected cumulative demand and resource constraints. Note that by taking expected cumulative demand, all time dynamics of the demand forecast are ignored in this formulation. Subsequently bid-prices are generated by solving the corresponding dual problem, where the optimal dual prices associated with the resource constraints are used. Given its simplicity and ease of implementation, the static linear programming approximation approach has been widely used in the practice, particularly for airline revenue management. The linear programs are resolved frequently over time to account for the static nature of the approximation, and to partially restore the time and capacity dependency of the bid-price. More recently, Adelman [1] proposes a new dynamic approximation method that tackles bid-price control directly in the dynamic setting. Adelman reformulates the dynamic programming equation as a gigantic linear program with exponential number of variables and constraints. He then proposes to use affine function to approximate the value function (where the coefficients of the affine function resemble bid-prices), and thus limit the the number of decision variables and make the problem tractable. Adelman show that the approximation with affine function can be solved efficiently with constraint generation, and more importantly provides a better theoretical bound to the true value function than that offered by the static linear programming approximation. Furthermore, the dynamic approximation also tends to provide better control policy in his experimental studies.

In this chapter we will show that both the static linear programming approximation and dynamic affine value function approximation can be easily adapted to the decentralized setting, and therefore their computational efficiency be exploited.

3.2 Static Linear Programming Approximation (LP)

3.2.1 Approximate Centralized Control

We first review the approximation method in centralized setting. Ignoring all the time dynamics, we can model the resource allocation problem as a simple one-stage planning problem:

$$(C^{LP}) \begin{cases} Z^{LP} \equiv \max_{Y} \sum_{j \in \mathcal{J}} r_j Y_j \\ \text{subject to:} \\ 0 \leq Y_j \leq \sum_{t=1}^{T} q_{jt} \quad \forall j, \\ \sum_{j \in J} a_{lj} Y_j \leq x_{l1} \quad \forall l, \end{cases} \tag{3.2}$$

where Y_j is the expected amount of bundle j we would like to sell over the entire horizon, and x_{l1} is the starting inventory for resource l at time 1. The objective of the one-stage planning problem is simply to maximize the expected revenue subject to expected cumulative demand and resource constraints. The corresponding dual problem is

$$(D^{LP}) \begin{cases} \min_{\pi,\theta} \sum_{l=1}^{K} \pi_l x_{l1} + \sum_{j \in \mathcal{J}} \theta_j \sum_{t=1}^{T} q_{jt} \\ \text{subject to:} \\ \sum_{l=1}^{K} \pi_l a_{lj} + \theta_j \geq r_j \quad \forall j, \\ \pi, \theta \geq 0, \end{cases} \tag{3.3}$$

where π_l is dual/shadow price for resource l, and will be used as the bid-price.

Let $\{\pi_l^*, \theta_j^*\}$ denote the optimal dual prices, the bid-price control policy following (3.1) is

$$\mu_j^*(x) = \begin{cases} 1 & \text{if } r_j \geq \sum_{l=1}^{K} \pi_l^* a_{lj}, \text{and } x \geq A_j, \\ 0 & \text{otherwise,} \end{cases} \tag{3.4}$$

where we accept the request for bundle j only if the revenue is greater than the sum of the bid-price of the resource required and there is enough inventory to satisfy the demand. Note that here π_l^* is static, but we can re-resolve (3.3) frequently overtime to get the most up-to-date bid prices and partially account for the time dynamics.

An important theoretical result shows that that the static linear programming approximation yields a upper bound to the optimal value function (Cooper [14]),

$$Z^{LP} \geq V(1, x_1), \tag{3.5}$$

and static policy is asymptotically optimal when the demand, capacity, and time horizon scales linearly, namely Z^{LP} converges to $V(1, x_1)$. Next we show that this approximation method can be adapted to the decentralized setting easily.

3.2.2 Approximate Decentralized Control

The static linear programming approximation of the centralized problem (3.2)-(3.3) can be decentralized in a natural way by inventory splitting. Let $x_{l1} = x_{l1}^{(1)} + \cdots + x_{l1}^{(m)}$, where we split the inventory among the agents, we can rewrite (3.2)-(3.3) as the followings to see the sharing structure more succinctly

$$(C^{LP}) \begin{cases} Z^{LP} \equiv \max_{Y^{(1)},\cdots,Y^{(m)}} \sum_{j \in \mathcal{J}_1} r_j Y_j^{(1)} + \cdots + \sum_{j \in \mathcal{J}_m} r_j Y_j^{(m)} \\ \text{subject to:} \\ 0 \leq Y_j^{(i)} \leq \sum_{t=1}^{T} q_{jt} \quad \forall j \in \mathcal{J}_i, i \in \mathcal{I}, \\ \sum_{j \in J_1} a_{lj} Y_j^{(1)} + \cdots + \sum_{j \in J_m} a_{lj} Y_j^{(m)} \leq x_{l1}^{(1)} + \cdots + x_{l1}^{(m)} \quad \forall l, \end{cases} \tag{3.6}$$

where the objective is the sum of the revenues generated by individual agents, and the resource constraint is where the coupling of the agents occurs. The corresponding dual problem can be written as

$$(D^{LP}) \begin{cases} \min_{\pi,\theta} \sum_{l=1}^{K} \pi_l (x_{l1}^{(1)} + \cdots + x_{l1}^{(m)}) + \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}_i} \theta_j^{(i)} \sum_{t=1}^{T} q_{jt} \\ \text{subject to:} \\ \sum_{l=1}^{K} \pi_l a_{lj} + \theta_j^{(i)} \geq r_j \quad \forall j \in \mathcal{J}_i, i \in I, \\ \pi, \theta \geq 0. \end{cases} \tag{3.7}$$

A closer look at (3.6)-(3.7) reveals that the formulation can be decentralized as the following,

$$(C_i^{LP}) \begin{cases} Z_i^{LP} \equiv \max_{Y^{(i)}} \sum_{j \in \mathcal{J}_i} r_j Y_j^{(i)} \\ \text{subject to:} \\ 0 \leq Y_j^{(i)} \leq \sum_{t=1}^{T} q_{jt} \quad \forall j \in \mathcal{J}_i, \\ \sum_{j \in J_i} a_{lj} Y_j^{(i)} \leq x_{l1}^{(i)}, \end{cases} \tag{3.8}$$

where Agent i maximizes his expected revenue subject to his allocated share of the total inventory $x_{l1}^{(i)}$. The corresponding dual problem is

$$(D_i^{LP})\begin{cases}\min\limits_{\pi^{(i)},\theta^{(i)}} \sum\limits_{l=1}^{K} \pi_l^{(i)} x_{l1}^{(i)} + \sum\limits_{j\in\mathcal{J}_i} \theta_j^{(i)} \sum\limits_{t=1}^{T} q_{jt} \\ \text{subject to:} \\ \sum\limits_{l=1}^{K} \pi_l^{(i)} a_{lj} + \theta_j^{(i)} \geq r_j \quad \forall j \in \mathcal{J}_i, \\ \pi^{(i)}, \theta^{(i)} \geq 0, \end{cases} \tag{3.9}$$

where Agent i can find the bid-price $\pi_l^{(i)}$ under his own demand forecast and the allocated inventory $x_{l1}^{(i)}$. Once the inventory is split among the agents, each becomes independent and can operate on his own. The key question is that how the inventory should be split such that no efficiency is lost by decentralizing the decision making to local agents.

A sufficient condition for no efficiency loss is such that over the same resource, all agents have the same bid-price as the central agent, i.e. $\pi_l^{(1)} = \cdots = \pi_l^{(m)} = \pi_l$. In such case, applying the bid-price control at local level is equivalent to doing so at the central level. Therefore, a simple strategy is to split inventory among agents, such that at equilibrium all agents' bid-price over the same resource converges. For example, suppose there are two agents, and over some resource l, Agent 1 has a bid-price of \$100, whereas Agent 2 has a higher bid-price of \$200, then it is better off for the overall system to allocate more inventory from Agent 1 to Agent 2, until their bid-price converge to the same value. If we can find such a allocation, will the bid-price converges to the one that's centrally optimal? The answer is yes, and we shall state this result in the following proposition.

Proposition 3.2.1. *If there exists an inventory allocation scheme,* $\left\{x_{l1}^{(i)*} | x_{l1}^{(1)*} + \cdots + x_{l1}^{(m)*} = x_{l1}, \forall l\right\}$*, such that at equilibrium the bid-price of all agents over the same resource converges,*

$$\pi_l^{(1)*} = \cdots = \pi_l^{(m)*} \qquad \forall l,$$

then the bid-price will also be equal to the centrally optimal bid-price,

$$\pi_l^{(1)*} = \cdots = \pi_l^{(m)*} = \pi_l^* \qquad \forall l,$$

where π_l^* solves (3.3). Therefore under such allocation scheme, each agent's decentralized pricing policy will also maximizes the system's net revenue.

Proof. We prove the claim by showing that at equilibrium, the optimal bid-prices for the decentralized problems are also optimal for the centralized problem, by satisfying primal and dual feasibility, as well as the complementary slackness conditions.

1. *Primal and Dual feasibility*:
   Clearly, any feasible solution set $\{Y_j^{(i)}, \pi_l^{(i)}, \theta_j^{(i)}, \forall i, j, l\}$ obtained from the decentralized problems (3.8)-(3.9) are also feasible for the centralized problem (3.6)-(3.7).

2. *Complementary slackness*:
Suppose $\{Y_j^{(i)*}, \pi_l^{(i)*}, \theta_j^{(i)*}, \forall i, j, l\}$ are optimal for (3.8)-(3.9), and we have $\pi_l^{(1)*} = \cdots = \pi_l^{(m)*} = \pi_l^*$, under some equilibrium allocation scheme $x_{l1}^{(1)*} + \cdots + x_{l1}^{(m)*} = x_{l1}$, then we have,
$$\begin{cases} \pi_l^{(i)*}\Big(x_{l1}^{(i)*} - \sum_{j\in\mathcal{J}_i} a_{lj} Y_j^{(i)*}\Big) = 0 \quad \forall i \in \mathcal{I}, l, \\ \pi_l^{(1)*} = \cdots = \pi_l^{(m)*} = \pi_l^*. \end{cases}$$
$$\Rightarrow \pi_l^*\Big(x_{l1}^{(1)*} + \cdots + x_{l1}^{(m)*} - \sum_{j\in\mathcal{J}_1} a_{lj} Y_j^{(1)*} - \cdots - \sum_{j\in\mathcal{J}_m} a_{lj} Y_j^{(m)*}\Big) = 0,$$
$$\Rightarrow \pi_l^*\Big(x_{l1} - \sum_{i\in I}\sum_{j\in\mathcal{J}_i} a_{lj} Y_j^{(i)*}\Big) = 0,$$
together with,
$$Y_j^{(i)*}\Big(\sum_{l=1}^{K} a_{lj}\pi_l^* + \theta_j^{(i)*} - r_j\Big) = 0 \quad \forall j, i,$$
we have shown that $\{Y_j^{(i)*}, \pi_l^{(i)*}, \theta_j^{(i)*}, \forall i, j, l\}$ also satisfy the complementary slackness conditions for the centralized problem.

□

Note that in (3.8)-(3.9) once the inventory is spit, individual agents' problems become independent. They will solve for their own bid-price, and need not worry about other agents' models and not even how many other agents exist in the system. Such features are identical to those in the decentralized framework we proposed in Chapter 2. Furthermore, we can easily adapt our iterative algorithm to solve for the optimal inventory allocation in decentralized fashion.

Algorithm 3.2.1. *(Approximate Iterative Decentralized Algorithm: LP)*

Initialize: Set k = 1, and choose some initial inventory allocation $x_{l1}^{(1),1} + \cdots + x_{l1}^{(m),1} = x_{l1}$.

Step 1: Given k, and $x_{l1}^{(i),k}$

- *Each agent solves (3.9) and computes his bid-price* $\pi_l^{(i),k}$;
- *Stop if, a satisfactory level of precision has been reached,*
$$\sup_{i\neq j} |\pi_l^{(i),k} - \pi_l^{(j),k}| \le \epsilon;$$
otherwise, each agent communicates the updated bid-price $\pi_l^{(i),k}$ *to the system.*
- *System reallocates the inventory, i.e. increase inventory for agents with higher bid-price, and decrease inventory for agents with lower bid-price.*

Step 2: Set k *to* $k+1$ *and return to Step 1.*

Note that in this approximate iterative decentralized algorithm, all is communicated is the updated bid-price, local agents need not reveal any of his sensitive private information. In summary, we have shown the static linear programming approximation of the centralized problem (3.6)-(3.7) can be adapted naturally into the decentralized setting, and lead to decentralized formulation for the local agent (3.8)-(3.9). This approach has several attractive features:

- Locally each agent only need to solve a linear program (i.e. the dual problem (3.9)), which can be solved efficiently. In practice, this approximation has been used widely in centralized setting to construct bid-price control policy.
- Decentralized control achieves central optimality under the static approximation framework, such that local bid-price converges to the centrally optimal bid-price, i.e. no efficiency loss within the static approximation framework.
- The static approximation generates control policy that is asymptotically optimal when the system parameters scale linearly.

However, the static approximation method is not without pitfalls. In particular the optimality with reference to the optimal dynamic control policy is only guaranteed asymptotically, which may not be practical for real-life applications. More generally, static approximation only provides an upper bound to the optimal value function and correspondingly a sub-optimal control policy. The main weakness of the approximation is its static nature, as time dynamics is totally ignored in the formulation (3.2). However, dynamic control is likely to be critical for real-life applications. For example, demands often change quickly over time, and exhibit patterns such as seasonality, peak and off-peak periods, and etc. Next we discuss a more dynamic approximation method, the affine value function approximation proposed by Adelman [1]. Adelman's approach is capable of producing a time-trajectory of the bid-price, which leads to tighter bound for the value function than that offered by the static linear programming approximation. Furthermore, Adelman's approach also produce better control policy as evidenced in his experimental studies.

3.3 Affine Value Function Approximation (AVF)

3.3.1 Approximate Centralized Control

We first review the approximation method in centralized setting. The idea begins with the observation that the optimal value function for the centralize problem at the initial value can be computed by a linear program (Adelman [1]),

$$(C)\begin{cases} \min\limits_{V(\cdot)} V(1,x_1) \\ \text{subject to:} \\ V(t,x) \geq \sum\limits_{i\in\mathcal{J}} q_{jt}\mu_{jt}(x)(r_j - \Delta V(t+1,x,A_j)) + V(t+1,x) \quad \forall t,x,\mu, \\ V(T+1,x) = 0. \end{cases} \tag{3.10}$$

However, it is impractical trying to solve (3.10) directly because it has exponential number of decision variables and constraints. Adelman proposes to approximate the dynamic value function with affine functions. Let $V(t,x)$ be the optimal value function solving (2.4), we can approximate $V(t,x)$ as

$$V(t,x) \approx \widehat{V}(t,x) = \theta_t + \sum_{l=1}^{K} v_{lt} x_l. \tag{3.11}$$

Replacing the optimal value function $V(t,x)$ with the affine functional approximation, we can approximate (3.10) with

$$(C^{AVF}) \begin{cases} Z^{AVF} \equiv \min_{\theta,v} \theta_1 + \sum_{l=1}^{K} v_{l1} x_{l1} \\ \text{subject to:} \\ \theta_t + \sum_{l=1}^{K} v_{lt} x_l \geq \sum_{i \in \mathcal{J}} q_{jt} \mu_{jt}(x)(r_j - \sum_{l=1}^{K} v_{l,t+1} a_{lj}) \\ \quad +\theta_{t+1} + \sum_{l=1}^{K} v_{l,t+1} x_l \quad \forall t, x, \mu, \\ \theta_{T+1}, v_{l,T+1} = 0. \end{cases} \tag{3.12}$$

This formulation reduces the number of variables drastically, and can be solved rather efficiently with constraint generation (Adelman [1]). Finally let $\{v_{lt}^*, \theta_t^*\}$ denote the optimal values of (3.12), the solutions fall nicely into the the bid-price control framework, with the policy given as

$$\mu_{jt}^*(x) = \begin{cases} 1 & \text{if } r_j \geq \sum_{l=1}^{K} v_{l,t+1}^* a_{lj}, \text{and } x \geq A_j, \\ 0 & \text{otherwise,} \end{cases} \tag{3.13}$$

The key difference between the static and dynamic approximation is that (3.12) computes a time-trajectory of the bid-price $\{v_{lt}^*, \forall l, t\}$, whereas (3.2) only compute static bid price using a one-stage planning approach. It can be shown that this dynamic approximation provides a tighter upper bound than that offered by static approximation:

$$Z^{LP} \geq Z^{AVF} \geq V(1, x_1). \tag{3.14}$$

Furthermore, through a number of numerical studies, Adelman demonstrates that the dynamic approximation tends to yield better bid-price control policies than that offered by the static approximation even with frequent resolving.

3.3.2 Approximate Decentralized Control

The affine value function approximation can be easily adapted to the decentralized setting. We apply directly the idea of replacing Agent i's value function with affine functional approximation,

$$(C_i^{AVF})\begin{cases} Z_i^{AVF} \equiv \min_{\theta^{(i)},v^{(i)}} \theta_1^{(i)} + \sum_{l=1}^{K} v_{l1}^{(i)}(R)x_{l1} \\ \text{subject to:} \\ \theta_t^{(i)} + \sum_{l=1}^{K} v_{lt}^{(i)}(R)x_l \geq \sum_{j\in\mathcal{J}_i} q_{jt}\mu_{jt}^{(i)}(x)(r_j - R_{-i}(t,x,A_j) - \sum_{l=1}^{K} v_{l,t+1}^{(i)}(R)a_{lj}) \\ \sum_{j\in\mathcal{J}_{-i}} \eta_{jt}^{(i)}(x)\Big[R_i(t,x,A_j) - \sum_{l=1}^{K} v_{l,t+1}^{(i)}(R)a_{lj}\Big] + \theta_{t+1}^{(i)} + \sum_{l=1}^{K} v_{l,t+1}^{(i)}(R)x_l \quad \forall t,x,\mu, \\ \theta_{T+1}^{(i)}, v_{l,T+1}^{(i)}(R) = 0, \end{cases} \tag{3.15}$$

where again we emphasize that Agent i's model $\eta_{jt}^{(i)}(x)$ for the probability that resource A_j are consumed by other agents may be mis-specified; also $v_{lt}^{(i)}(R)$ makes explicit that Agent i's value function depends on the contract chosen R.

Parallel to the development in Chapter 2, a sensible choice of contract R should be defined by the agent's marginal valuation of the inventory, and can be constructed in similar way as (2.14)-(2.15). Let $\{v_{lt}\}$ be the set of bid-price of the centralized Agent (3.12), we can solve for Agent i's bid-price approximately by solving the following

$$\begin{cases} \{\hat{\theta}^{(i)}, \hat{v}^{(i)}\} \equiv \text{argmin } \theta_1^{(i)} + \sum_{l=1}^{K} v_{l1}^{(i)} x_{l1} \\ \text{subject to:} \\ \theta_t^{(i)} + \sum_{l=1}^{K} v_{lt}^{(i)} x_l \geq \sum_{j\in\mathcal{J}_i} q_{jt}\mu_{jt}^{(i)}(x)(r_j - \sum_{l=1}^{K} v_{l,t+1}a_{lj}) \\ \quad +\theta_{t+1}^{(i)} + \sum_{l=1}^{K} v_{l,t+1}^{(i)}(R)x_l \quad \forall t,x,\mu, \\ \theta_{T+1}^{(i)}, v_{l,T+1}^{(i)} = 0. \end{cases} \tag{3.16}$$

Finally the contract R can be approximately constructed as

$$\begin{cases} \hat{R}_{-i}(t,x,A_j) &= \sum_{i'\neq i}\sum_{l=1}^{K} \hat{v}_{l,t+1}^{(i')} a_{lj}, \\ \hat{R}_i(t,x,A_j) &= \sum_{l=1}^{K} \hat{v}_{l,t+1}^{(i)} a_{lj}. \end{cases} \tag{3.17}$$

Given $\{\hat{v}_{lt}^{(i)}, \hat{R}_{-i}, \hat{R}_i\}$, Agent i's local bid-price policy can be defined as,

$$\mu_{jt}^*(x) = \begin{cases} 1 & \text{if } r_j \geq \sum_{l=1}^{K} \hat{v}_{l,t+1}^{(i)} a_{lj} + \hat{R}_{-i}(t,x,A_j) = \sum_{i\in\mathcal{I}} \sum_{l=1}^{K} \hat{v}_{l,t+1}^{(i)} a_{lj}, \\ 0 & \text{otherwise,} \end{cases} \tag{3.18}$$

Therefore, the role played by the contract $\hat{R}$ is to ensure local Agent account for other agents' valuation of the resource that is being consumed, and we are essentially approximating central Agent's bid-price by the sum of local agents' approximate bid prices given by (3.16),

$$v_{lt} \approx \sum_{i\in\mathcal{I}} \hat{v}_{lt}^{(i)}. \tag{3.19}$$

Proposition 3.3.1. *With contract $\hat{R}$ given by (3.16)-(3.17), it is easy to show that $\sum_{i\in\mathcal{I}} \hat{v}_{lt}^{(i)}$ is a feasible solution to the centralized problem (3.12), therefore offering a upper bound to central Agent's approximate problem*

$$\sum_{i\in\mathcal{I}} Z_i^{AVF} \geq Z^{AVF} \geq V(1,x_1). \tag{3.20}$$

Note that in (3.16)-(3.17) the approximate $\hat{R}$ is constructed assuming the set of bid-price $\{v_{lt}\}$ of the centralized Agent is given, which is unrealistic and defeating the purpose of decentralization. Fortunately, the local approximate bid-price $\{\hat{v}_{lt}^{(i)}\}$ and contracts $\{\hat{R}_{-i}, \hat{R}\}$ can be constructed in decentralized manner using our iterative algorithm:

Algorithm 3.3.1. *(Approximate Iterative Decentralized Algorithm: AVF)*

Initialize: Set $k = 1$, and $\hat{R}_i^1(t,x,A_j) = 0, \hat{R}_{-i}^1(t,x,A_j) = 0$.

Step 1: Given k, and $\hat{R}_i^k(t,x,A_j), \hat{R}_{-i}^k(t,x,A_j)$

- *Each agent solves his approximate decentralized problem (3.15) and computes his approximate bid-price $\hat{v}_{lt}^{(i),k}$.*
- *Stop if, a satisfactory level of precision has been reached,*

$$\sup_{i,t,l} |\hat{v}_{lt}^{(i),k+1} - \hat{v}_{lt}^{(i),k}| \leq \epsilon;$$

otherwise, each agent updates his approximate transfer contract,

$$\hat{R}_i^{k+1}(t,x,A_j) = \sum_{l=1}^{K} \hat{v}_{l,t+1}^{(i),k} a_{lj}.$$

- *Each agent communicates the updated approximate transfer contract $\hat{R}_i^{k+1}(t,x,A_j)$ to the system.*

- *System synthesize* $\hat{R}_{-i}^{k+1}(t,x,A_j) = \sum_{i' \neq i} \hat{R}_i^{k+1}(t,x,A_j)$*, and broadcast them back to all agents.*

Step 2: Set k *to* $k+1$ *and return to Step 1.*

In summary we have shown the affine value function approximation of the centralized problem (3.12) can be adapted easily into the decentralized setting, leading to decentralized formulation for the local agent (3.15). The decentralization strategy requires the selling agent to transfer part of the revenue to other agents as compensation whenever a request is accepted and system resource gets consumed. Note that this approximation method inherit the same structure of the coordination mechanism we propose in Chapter 2, as opposed to the static linear programming approximation where an inventory splitting approach is used. Moreover, the affine value function approximation is a more dynamic approach, and computes a time-trajectory of the bid-price. We have shown that by introducing transfer contracts, local bid-price can be constructed to approximate the optimal bid-price that could be obtained if there exist a centralized agent. Further, the bid-price and the associated contracts can be constructed iteratively without requiring local agent to reveal private information. Therefore, we have inherited all of the operational features we proposed in Chapter 2. The affine value function approximation is an promising approach as Adelman has demonstrated that in centralized setting, the bid-price policies tend to out-perform those offered by the static linear programming approximation. We hope to carry out numerical studies to evaluate the dynamic approximation in decentralized setting in our future work.

3.4 Summary

We discussed one static and one dynamic approximation methods in this chapter. The approximation methods are originally developed for centralized problems, but we show that they can be easily adapted to decentralized setting. For the static approximation using an one-stage linear programming formulation, we show that a simple inventory splitting approach can help decentralize the decision making, and the locally optimal bid-prices at equilibrium will converges to the centrally optimal bid-prices. However, the weakness of this approach is that the bid-prices computed are static in nature, and the corresponding control policy could be far from optimal. For the dynamic approximation using affine functional, we show that the approximation methods can be adapted seamlessly into our decentralized framework. The approximation basically reduces our solution space to parameterized affine functionals, and we can still construct contracts and let each agents evaluate their value functions in a decentralized fashion iteratively. Finally, the various bounds of the two approximation methods are summarized as follows:

$$\begin{cases} \text{Centralized:} & Z^{LP} \geq Z^{AVF} \geq V(1,x_1), \\ \text{LP:} & \sum_{i \in \mathcal{I}} Z_i^{LP} = Z^{LP} \geq V(1,x_1), \\ \text{AVF:} & \sum_{i \in \mathcal{I}} Z_i^{AVF} \geq Z^{AVF} \geq V(1,x_1). \end{cases} \tag{3.21}$$

Chapter 4

Decentralized Portfolio and Risk Management

In this chapter we formulate a class of related decentralized stochastic dynamic control problems but specialize to portfolio and risk management. We consider a financial firm that trade in multiple correlated markets (e.g. fixed income, equity, commodities, real estate, and etc). The system is in general too complex to be managed by a single centralized agent, and multiple agents (portfolio managers) are typically hired to manage investment decisions in the different markets. Given certain risk capitals, each agent solves his own portfolio selection problem and seeks to optimize his own risk and return profile. The firm's investment performance depends on the net holdings of all agents' individual holdings, and its optimal risk and return profile can be quite different from those of the individual agents. Decentralized investment management is necessary here but the system's efficiency loss due to the loss of diversification when agents are left uncoordinated can be substantial. In this chapter we seek to investigate how to set up proper incentive structure so as to induce decentralized agents to select portfolios that in the aggregate will optimize the risk and return profile for the overall firm.

This chapter is organized as follows. Section 4.1 describes the investment model for markets, returns, risk factors, and the firm's utility function. Section 4.2 and 4.3 formulate both a centralized as well as decentralized approach for making investment decisions for this system. Section 4.4 characterizes the optimal incentive structure in the form of swap contracts, and establish results about existence and optimality. Section 4.5 provides an iterative algorithm for constructing the set of optimal contracts in decentralized manner, and proves convergence. Finally Section 4.6 concludes this chapter.

4.1 General Description

We now describe the investment model for markets, returns, risk factors, and the firm's utility function. In subsequent sections, we formulate the centralized and decentralized approaches for making investment decisions for this system.

4.1.1 Market Model

Uncertainty is modeled by an $n+m$ dimensional standard Brownian motion

$$[W_1(t), \cdots, W_n(t), V_1(t), \cdots, V_m(t)].$$

There are $i = 1, \cdots, m$ markets. To ease notation, we assume that each market consists of a single risky asset whose price evolves according to the stochastic differential equation

$$dS_i(t) = \alpha_i S_i(t)dt + S_i(t)\Big\{\sum_{j=1}^{n} \sigma_{ij} dW_j(t) + \epsilon_i dV_i(t)\Big\}. \tag{4.1}$$

Our analysis extends with no essential difficulty to the multi-dimensional case. The n components $W_1(t), \cdots, W_n(t)$ are common factors and dependence between the assets $S_1(t), \cdots, S_m(t)$ comes about due to exposure to these common risk factors. The terms $V_1(t), \cdots, V_m(t)$ model idiosyncratic risk for the assets in each of the m markets. For example, in the case of the asset $S_i(t)$ in market i, $\{\sigma_{i1}, \cdots, \sigma_{in}\}$ are its exposures to the n common risk factors and ϵ_i corresponds to its idiosyncratic risk.

Note that in (4.1) the common risk factors are independent. First this is without loss of generality. Second it offers a natural and disciplined way of attributing risks into the different risk factors, and later will be shown to be very nice property that offers investor the flexibility of trading off optimality for operational efficiency using our proposed coordination mechanism. Now consider a somewhat more general factor model (which we will show to reduce to (4.1)), where asset returns can be explained by n correlated factors $I_1(t), \cdots, I_n(t)$

$$dS_i(t) = \widetilde{\alpha}_i S_i(t)dt + S_i(t)\Big\{\sum_{k=1}^{n} \beta_{ik} dI_k(t) + \epsilon_i dV_i(t)\Big\}. \tag{4.2}$$

The factor $I_k(t)$ can be modeled as

$$dI_k(t) = \delta_k dt + dB_k(t), \quad k = 1, ..., n, \tag{4.3}$$

where we assume the standard Brownian motion $B_1(t), ..., B_n(t)$ are correlated

$$\begin{aligned} Cov(dB_k(t), dB_k(t)) &= dt, \quad \forall k, \\ Cov(dB_k(t), dB_l(t)) &= \rho_{kl} dt, \quad \forall k \neq l. \end{aligned}$$

Let $\rho = [\rho_{kl}]$ denotes the covariance matrix. It is well known that there exist an orthonormal basis such that the covariance matrix of $dB_1(t), \cdots, dB_n(t)$ expressed in this new basis is diagonal,

$$\begin{aligned} dW(t) &= \Lambda^{-\frac{1}{2}} Q' dB(t), \\ Cov(dW(t)) &= \mathbb{I}, \end{aligned}$$

where Q is the orthonormal basis whose axes (columns) are called the principal components of ρ, and Λ is a diagonal matrix whose diagonal consists of the eigenvalues of ρ. The new variables $dW_1(t), \cdots, dW_n(t)$ are projections of $dB_1(t), \cdots, dB_n(t)$ onto its principal components, and are uncorrelated, and with proper scaling by the eigenvalues $\Lambda^{-\frac{1}{2}}$ has covariance matrix equal to the identify matrix $\mathbb{I}$.

Working in the new basis has the advantage that the new risk factors are now uncorrelated, and risks are additive and can be easily attributed into individual factors. This fact motivates us to set up the model in terms of uncorrelated risk factors, which later would allow us to be able to prioritize implementation of the swap contract among the different risk factors according to their contribution to the firm's overall risk (more to follow in the next chapter).

Now we show that we can re-write (4.2) in terms of the set of independent risk factors $dW(t)$, it is easy to see that we have

$$dB(t) = Q\Lambda^{\frac{1}{2}}dW(t) = \sum_{j=1}^{n} Q^{(j)}\Lambda_j^{\frac{1}{2}}dW_j(t), \tag{4.4}$$

where $W_j(t)$ is the j^{th} new risk factor by projecting $B_1(t), \cdots, B_n(t)$ onto its j^{th} principal component, where $Q^{(j)}$ is the j^{th} principal component, and $\Lambda_j^{\frac{1}{2}}$ is the corresponding eigenvalue. Finally substituting (4.3) and (4.4) into (4.2), we have

$$\begin{aligned}
dS_i(t) &= \widetilde{\alpha}_i S_i(t)dt + S_i(t)\Big\{\sum_{k=1}^{n}\beta_{ik}[\delta_k dt + dB_k(t)] + \epsilon_i dV_i(t)\Big\}, \\
&= [\widetilde{\alpha}_i + \sum_{k=1}^{n}\beta_{ik}\delta_k]S_i(t)dt + S_i(t)\Big\{\sum_{k=1}^{n}\beta_{ik}[\sum_{j=1}^{n}Q_{kj}\Lambda_j^{\frac{1}{2}}dW_j(t)] + \epsilon_i dV_i(t)\Big\}, \\
&= [\widetilde{\alpha}_i + \sum_{k=1}^{n}\beta_{ik}\delta_k]S_i(t)dt + S_i(t)\Big\{\sum_{j=1}^{n}[\sum_{k=1}^{n}\beta_{ik}Q_{kj}]\Lambda_j^{\frac{1}{2}}dW_j(t) + \epsilon_i dV_i(t)\Big\}, \\
&= \alpha_i S_i(t)dt + S_i(t)\Big\{\sum_{j=1}^{n}\sigma_{ij}dW_j(t) + \epsilon_i dV_i(t)\Big\},
\end{aligned}$$

which reduces to (4.1) with

$$\alpha_i = \widetilde{\alpha}_i + \sum_{k=1}^{n}\beta_{ik}\delta_k,$$

$$\text{and } \sigma_{ij} = [\sum_{k=1}^{n}\beta_{ik}Q_{kj}]\Lambda_j^{\frac{1}{2}},$$

where σ_{ij} is precisely the risk exposure to the i^{th} market on the j^{th} new and uncorrelated risk factor $W_j(t)$.

In summary, we have shown that we can begin with a general factor model such that the underlying factors are correlated, and by projecting the risk factors onto its principal components we can re-express the firm's risk exposures in terms of a set of uncorrelated risk factors (4.1). This representation is not only mathematically neat and convenient, but more importantly it allows us to generate insight as how the firm's total risk can be attributed into different independent sources, and how contracts on different factors can be prioritized such that we need only focus on a small subset of the critical risk factors to generate the most improvement to the overall portfolio performance.

4.1.2 Investment Model

The firm has holdings in each market. Let $x(t)$ denote the firm's net wealth and $\pi_i(t)$ the dollar value of its portfolio in market i. Then

$$dx(t) = \sum_{i=1}^{m} \pi_i(t)\alpha_i dt + \sum_{i=1}^{m} \pi_i(t)\Big\{ \sum_{j=1}^{n} \sigma_{ij} dW_j(t) + \epsilon_i dV_i(t)\Big\}. \tag{4.5}$$

We assume that the firm evaluates investment policy $\pi(t)$ using an generalized exponential utility function. Before introducing the generalized exponential utility, we shall first discuss the standard exponential utility function $\mathbb{E}[-exp\{-x(T)/\theta\}]$. Clearly maximizing the exponential utility is equivalent to maximizing its log-transform as defined below.

Definition 4.1.1. *Standard (log-transformed) exponential utility*

$$\begin{aligned} & -\theta \ln \mathbb{E}_{\mathbb{P}}\Big[e^{-\frac{1}{\theta}x(T)}\Big] \\ = \;& \mathbb{E}_{\mathbb{P}}\Big[x(T)\Big] - \theta \ln \mathbb{E}_{\mathbb{P}}\Big[e^{-\frac{1}{\theta}(x(T)-\mathbb{E}_{\mathbb{P}}[x(T)])}\Big], \\ = \;& \mathbb{E}_{\mathbb{P}}\Big[x(T)\Big] - \mathcal{R}_\theta(x(T)), \end{aligned} \tag{4.6}$$

where

$$\mathcal{R}_\theta(x(T)) \equiv \theta \ln \mathbb{E}_{\mathbb{P}}\Big[e^{-\frac{1}{\theta}(x(T)-\mathbb{E}_{\mathbb{P}}[x(T)])}\Big], \tag{4.7}$$

is a so-called convex risk measure [16] commonly referred to as the exponential premium. This allows us to interpret exponential utility as a risk-adjusted expected return.

Exponential utility has a well-known dual representation which we define next.

Definition 4.1.2. *Standard (log-transformed) exponential utility in dual represenation*

$$\begin{aligned} -\theta \ln \mathbb{E}_{\mathbb{P}}\Big[e^{-\frac{1}{\theta}x(T)}\Big] \;&\equiv\; \mathbb{E}_{\mathbb{P}}\Big[x(T)\Big] - \mathcal{R}_\theta(x(T)), \\ &=\; \min_{\gamma,\widetilde{\gamma}} \; \mathbb{E}_{\mathbb{Q}}\Big\{x(T) + \frac{\theta}{2}\sum_{j=1}^{n}\int_0^T |\gamma_j(t)|^2 dt \\ &\quad + \frac{\theta}{2}\sum_{i=1}^{m}\int_0^T |\widetilde{\gamma}_i(t)|^2 dt\Big\}. \end{aligned} \tag{4.8}$$

where under the new measure $\mathbb{Q}$ the return dynamic is

$$\begin{aligned} dx(t) &= \sum_{i=1}^{m} \pi_i(t) \Big\{ \alpha_i - \sum_{j=1}^{n} \gamma_j(t)\sigma_{ij} - \widetilde{\gamma}_i(t)\epsilon_i \Big\} dt \\ & \quad + \sum_{i=1}^{m} \pi_i(t) \Big\{ \sum_{j=1}^{n} \sigma_{ij} d\widetilde{W}_j(t) + \epsilon_i d\widetilde{V}_i(t) \Big\}. \end{aligned} \tag{4.9}$$

Under the dual representation it is natural to generalize the exponential utility such that the firm can have different risk tolerances toward difference sources of risk, which lead to our next definition of the generalized exponential utility.

Definition 4.1.3. *Generalized (log-transformed) exponential utility in dual representation*

$$\begin{aligned} & \mathbb{E}_{\mathbb{P}}\Big[x(T)\Big] - \mathcal{R}_{(\theta,\widetilde{\theta})}(x(T)) \\ \equiv \; & \min_{\gamma,\widetilde{\gamma}} \; \mathbb{E}_{\mathbb{Q}}\Big\{ x(T) + \sum_{j=1}^{n} \frac{\theta_j}{2} \int_0^T |\gamma_j(t)|^2 dt + \sum_{i=1}^{m} \frac{\widetilde{\theta}_i}{2} \int_0^T |\widetilde{\gamma}_i(t)|^2 dt \Big\}, \end{aligned} \tag{4.10}$$

under which the firm can specify different level of risk-tolerances toward different sources of risks, more specifically

- $\{\theta_1, \cdots, \theta_n\}$ *are risk tolerances toward exposures of common risk factors* $\{W_1, \cdots, W_n\}$,
- $\{\widetilde{\theta}_1, \cdots, \widetilde{\theta}_m\}$ *are risk tolerances toward exposure of idiosyncratic risk factors* $\{V_1, ..., V_m\}$.

Finally the firm evaluates investment policy $\pi(t) = \{\pi_1(t), \cdots, \pi_m(t)\}$ under the generalized (log-transformed) exponential utility and has the following expected terminal utility

$$\left\{ \begin{aligned} & \mathbb{E}_{\mathbb{P}}\Big[x(T)\Big] - \mathcal{R}_{(\theta,\widetilde{\theta})}(x(T)) \\ & \equiv \min_{\gamma,\widetilde{\gamma}} \; \mathbb{E}_{\mathbb{Q}}\Big\{ x(T) + \sum_{j=1}^{n} \frac{\theta_j}{2} \int_0^T |\gamma_j(t)|^2 dt + \sum_{i=1}^{m} \frac{\widetilde{\theta}_i}{2} \int_0^T |\widetilde{\gamma}_i(t)|^2 dt \Big\} \\ & \text{subject to:} \\ & dx(t) = \sum_{i=1}^{m} \pi_i(t) \Big\{ \alpha_i - \sum_{j=1}^{n} \gamma_j(t)\sigma_{ij} - \widetilde{\gamma}_i(t)\epsilon_i \Big\} dt \\ & \qquad + \sum_{i=1}^{m} \pi_i(t) \Big\{ \sum_{j=1}^{n} \sigma_{ij} d\widetilde{W}_j(t) + \epsilon_i d\widetilde{V}_i(t) \Big\}, \\ & x(0) = x_0. \end{aligned} \right. \tag{4.11}$$

The key difference between the centralized and decentralized portfolio choice problem is the way in which the investment policy $\pi(t) = \{\pi_1(t), \cdots, \pi_m(t)\}$ is obtained. In the centralized case, it is obtained by optimizing (4.11) with complete knowledge of all return forecasts $\alpha = [\alpha_1, \cdots, \alpha_m]$ and the risk model $\sigma = [\sigma_{ij}]_{i,j}, \epsilon = [\epsilon_i]_i$ connecting all of the markets. In the decentralized case, investment policies for different markets (i.e. different components of $\pi(t)$) are computed by different agents where each maximizes his own utility with knowledge only of the dynamics of the assets in his own market.

4.2 Centralized Agent's Problem

In this section we formulate a centralized version of the investment problem (4.11) and characterize the value function and optimal policy. The results in this section serve as a benchmark for the decentralized model that we introduce in Section 4.3.

The centralized agent's objective is to maximize terminal expected utility within specified levels of risk-tolerance $\{\theta, \widetilde{\theta}\}$. He does so with complete knowledge of the return forecasts and risk models of all markets $\{\alpha, \sigma, \epsilon\}$:

$$\begin{cases} \max_{\pi} \mathbb{E}_{\mathbb{P}}\Big[x(T)\Big] - \mathcal{R}_{(\theta,\widetilde{\theta})}(x(T)) \\ \equiv \max_{\pi} \min_{\gamma,\widetilde{\gamma}} \ \mathbb{E}_{\mathbb{Q}}\Big\{ x(T) + \sum_{j=1}^{n} \frac{\theta_j}{2} \int_0^T |\gamma_j(t)|^2 dt + \sum_{i=1}^{m} \frac{\widetilde{\theta}_i}{2} \int_0^T |\widetilde{\gamma}_i(t)|^2 dt \Big\} \\ \text{subject to:} \\ dx(t) = \sum_{i=1}^{m} \pi_i(t) \Big\{ \alpha_i - \sum_{j=1}^{n} \gamma_j(t)\sigma_{ij} - \widetilde{\gamma}_i(t)\epsilon_i \Big\} dt \\ \qquad + \sum_{i=1}^{m} \pi_i(t) \Big\{ \sum_{j=1}^{n} \sigma_{ij} d\widetilde{W}_j(t) + \epsilon_i d\widetilde{V}_i(t) \Big\}, \\ x(0) = x_0. \end{cases} \tag{4.12}$$

It follows from dynamic programming that the value function is

$$\begin{cases} V(t,\, x) = x + H(t), \\ \text{where,} \\ 0 = \dot{H}(t) + \max_{\pi} \min_{\gamma,\widetilde{\gamma}} \sum_{i=1}^{m} \pi_i \Big\{ \alpha_i - \sum_{j=1}^{n} \gamma_j \sigma_{ij} - \widetilde{\gamma}_i \epsilon_i \Big\} \\ \qquad + \sum_{j=1}^{n} \frac{\theta_j}{2} |\gamma_j|^2 + \sum_{i=1}^{m} \frac{\widetilde{\theta}_i}{2} |\widetilde{\gamma}_j|^2, \\ H(T) = 0. \end{cases} \tag{4.13}$$

The optimality equation can be rewritten more succinctly in matrix-vector forms as

$$0 = \dot{H}(t) + \max_{\pi} \min_{\gamma,\widetilde{\gamma}} \pi' \Big\{ \alpha - \sigma \cdot \gamma - diag(\epsilon) \cdot \widetilde{\gamma} \Big\} + \frac{1}{2} \gamma' \cdot diag(\theta) \cdot \gamma + \frac{1}{2} \widetilde{\gamma}' \cdot diag(\widetilde{\theta}) \cdot \widetilde{\gamma},$$

First minimizing over $\gamma_j, \widetilde{\gamma}_j$ we have

$$\begin{aligned} \gamma^*(t) &= [diag(\theta)]^{-1} \sigma' \pi, \\ \widetilde{\gamma}^*(t) &= [diag(\widetilde{\theta})]^{-1} diag(\epsilon) \pi. \end{aligned}$$

Substituting back into the optimality equation, we have

$$0 = \dot{H}(t) + \max_{\pi} \Big\{ \pi'\alpha - \frac{1}{2}\pi'\sigma \cdot [diag(\theta)]^{-1} \cdot \sigma'\pi - \frac{1}{2}\pi'[diag(\epsilon) \cdot [diag(\widetilde{\theta})]^{-1} \cdot diag(\epsilon)]\pi \Big\}.$$

Therefore the value function is

$$\begin{cases} V(t, x) = x + H(t), \\ \text{where,} \\ 0 = \dot{H}(t) + \max_{\pi} \Big\{ \pi'\alpha - \frac{1}{2}\pi'\sigma \cdot [diag(\theta)]^{-1} \cdot \sigma'\pi \\ \qquad\qquad - \frac{1}{2}\pi'[diag(\epsilon) \cdot [diag(\widetilde{\theta})]^{-1} \cdot diag(\epsilon)]\pi \Big\}, \\ H(T) = 0, \end{cases} \tag{4.14}$$

where the optimal portfolio (Merton [26])

$$\pi^*(t) = (\sigma \cdot [diag(\theta)]^{-1} \cdot \sigma' + diag(\epsilon) \cdot [diag(\widetilde{\theta})]^{-1} \cdot diag(\epsilon))^{-1}\alpha, \tag{4.15}$$

is the maximizer in (4.14) which solves

$$\alpha - (\sigma \cdot [diag(\theta)]^{-1} \cdot \sigma' + diag(\epsilon) \cdot [diag(\widetilde{\theta})]^{-1} \cdot diag(\epsilon))\pi = 0. \tag{4.16}$$

In formulating problem (4.12), it is implicitly assumed that the centralized agent knows all relevant system parameters including return forecasts for each asset and correlations between them. In many applications, such a knowledgeable agent does not exist but rather, that there are many decentralized agents where each specializes in a particular market sector. In such situations, it is not possible to formulate the centralized problem (4.12). Finally, observe that the centralized agent's optimal portfolio (4.15) is static. This convenient property is a consequence of the generalized exponential utility function in (4.12) and simplifies the analysis of this problem. That said, the main results in this paper can be generalized to systems where the optimal control depends on the state, albeit with more complicated contracts (see for example [12, 13, 24]).

4.3 Decentralized Agent's Problem

Several elements distinguish the centralized and decentralized problems. Firstly, the decentralized system consists of multiple agents where each is responsible for the investment decisions in his own market. Secondly, every agent is better informed about the market he is investing in but less informed about those managed by others and is ignorant about the correlation structure between them. Finally, each agent maximizes his own objective function conditional on his information while the firm is concerned about the utility of net positions over all agents.

In this section, we formulate agent level dynamics and objectives and present an example showing that the efficiency loss of uncoordinated locally optimizing decentralized control can be substantial. A key insight from the example is that inefficiencies arise due to an incorrect pricing of risk, which leads us to introducing swap contracts between agents as a coordination mechanism. Swaps are derivative contracts which define cash transfer between agents. They enable risk transfer between agents and implicitly set a price of risk, which modifies agent level objectives objectives. We show that swap contracts can be found under which the decentralized agents achieve centralized optimality and provide an explicit characterization.

4.3.1 Agent i's Model

Agent i trades in market i with one risky asset

$$dS_i(t) = S_i(t)\alpha_i dt + S_i(t)\Big\{ \sum_{j=1}^{n} \sigma_{ij} dW_j(t) + \epsilon_i dV_i(t) \Big\}.$$

The return forecast α_i, and risk exposures $\sigma_{i1}, \cdots, \sigma_{in}, \epsilon_i$ are known by Agent i, though not by the others and the risk tolerance $\{\theta_{i1}, \cdots, \theta_{in}, \widetilde{\theta}_i\}$ is specified by the firm (though private to Agent i). We assume that the firm's total risk tolerance is the sum of those for individual agents $\theta = \sum_{i=1}^{m}\sum_{j=1}^{n} \theta_{ij} + \sum_{i=1}^{m} \widetilde{\theta}_i$.

Left to his own devices, Agent i solves

$$\begin{cases} \max_{\pi_i} \mathbb{E}_{\mathbb{P}}\Big[x_i(T)\Big] - \mathcal{R}_{(\theta_i,\widetilde{\theta}_i)}(x_i(T)) \\ \equiv \max_{\pi_i} \min_{\gamma,\widetilde{\gamma}} \ \mathbb{E}_{\mathbb{Q}}\Big\{ x_i(T) + \sum_{j=1}^{n} \frac{\theta_{ij}}{2} \int_0^T |\gamma_{ij}(t)|^2 dt + \frac{\widetilde{\theta}_i}{2} \int_0^T |\widetilde{\gamma}_i(t)|^2 dt \Big\} \\ \text{subject to:} \\ dx_i(t) = \pi_i(t)\Big\{ \alpha_i - \sum_{j=1}^{n} \gamma_{ij}(t)\sigma_{ij} - \widetilde{\gamma}_i(t)\epsilon_i \Big\} dt \\ \qquad\qquad + \pi_i(t) \Big\{ \sum_{j=1}^{n} \sigma_{ij} d\widetilde{W}_j(t) + \epsilon_i d\widetilde{V}_i(t) \Big\}, \\ x_i(0) \text{ given.} \end{cases} \tag{4.17}$$

Asset holdings $\{\pi_1(t), \cdots, \pi_m(t)\}$ in each of the m markets correspond to the optimal policies for each of the m decentralized agents.

4.3.2 A Running Example: Efficiency Loss

Consider a small portfolio consisting of 8 asset classes, the data (from Brennan et al. [10]) is based on (annualized) monthly historic statistics from January 1978 to December 1995. Table 4.1 and 4.2 shows the annualized asset returns, standard deviation, and correlation data. The variance-covariance matrix Σ_R can be decomposed into two components, namely the common factor risk component Σ_F and the idiosyncratic risk component Σ_ϵ:

$$\Sigma_R = \Sigma_F + \Sigma_\epsilon = \sigma\sigma' + diag(\epsilon)^2.$$

Table 4.3 shows the decomposition of the variance-covariance matrix, where the idiosyncratic part accounts for a typical 30% of the total variance. Under the framework of generalized exponential utility, we choose a profile of risk tolerance that differ by agents and the sources of risks, the numbers are in the typical range of 7.5 to 0.22, as shown in Table 4.4. The choice of this particular risk tolerance profile is somewhat arbitrary, we merely want to reflect that in reality the firm faces a heterogenous set of agents who can have different level of risk tolerances toward different sources of risk.

Asset Name	Asset No.	Return	Std. Dev.
Canadian Equity	1	11.64	19.05
French Equity	2	17.52	24.35
German Equity	3	13.32	21.55
Japanese Equity	4	17.52	24.39
U.K. Equity	5	16.44	20.82
U.S. Equity	6	15.48	14.90
U.S. Bonds	7	9.96	6.96
European Bonds	8	10.20	5.40

Table 4.1: Average historic return and standard deviation, all numbers in annualized percentage

	1	2	3	4	5	6	7	8
1	1.00	0.41	0.30	0.25	0.58	0.71	0.26	0.33
2		1.00	0.62	0.42	0.54	0.44	0.22	0.26
3			1.00	0.35	0.48	0.34	0.27	0.28
4				1.00	0.40	0.22	0.14	0.16
5					1.00	0.56	0.25	0.29
6						1.00	0.36	0.42
7							1.00	0.92
8								1.00

Table 4.2: Correlation matrix

Finally, after solving (4.12) and (4.17), Table 4.5 compares the optimal holdings chosen by the centralized agent and un-coordinated decentralized agents. Note the large difference in portfolio weights between the optimal holdings of the centralized agent relative to the those of the uncoordinated decentralized agents. The difference in risk-adjusted profit is also large: 8.64 versus 5.20, i.e. the firm loses around 40% in risk-adjusted profit going from centralized to un-coordinated decentralized setting.

There are several important lessons:

	Σ_F								Σ_ϵ
1	2.54	1.33	0.86	0.81	1.61	1.41	0.24	0.24	1.09
2	1.33	4.15	2.28	1.75	1.92	1.12	0.26	0.24	1.78
3	0.86	2.28	3.25	1.29	1.51	0.76	0.28	0.23	1.39
4	0.81	1.75	1.29	4.16	1.42	0.56	0.17	0.15	1.78
5	1.61	1.92	1.51	1.42	3.03	1.22	0.25	0.23	1.30
6	1.41	1.12	0.76	0.56	1.22	1.55	0.26	0.24	0.67
7	0.24	0.26	0.28	0.17	0.25	0.26	0.34	0.24	0.15
8	0.24	0.24	0.23	0.15	0.23	0.24	0.24	0.20	0.09

Table 4.3: Decomposition of the variance-covariance matrix into the factor component Σ_F and idiosyncratic component Σ_ϵ, all numbers in annualized percentage

	θ_{ij}								$\widetilde{\theta}_i$
1	4.78	3.19	2.39	1.91	1.59	1.37	1.20	1.06	7.50
2	3.19	2.13	1.59	1.28	1.06	0.91	0.80	0.71	5.00
3	2.39	1.59	1.20	0.96	0.80	0.68	0.60	0.53	3.75
4	1.91	1.28	0.96	0.77	0.64	0.55	0.48	0.43	3.00
5	1.59	1.06	0.80	0.64	0.53	0.46	0.40	0.35	2.50
6	1.37	0.91	0.68	0.55	0.46	0.39	0.34	0.30	2.14
7	1.20	0.80	0.60	0.48	0.40	0.34	0.30	0.27	1.88
8	1.06	0.71	0.53	0.43	0.35	0.30	0.27	0.24	1.67

Table 4.4: Risk tolerances towards different source of risk

Agent	Centrally Opt.	Locally Opt.
1	6.0%	23.9%
2	7.2%	12.7%
3	2.9%	7.9%
4	5.8%	5.4%
5	2.7%	5.2%
6	7.5%	7.6%
7	18.9%	15.1%
8	48.9%	22.3%
Risk-adjusted profit	8.64	5.20

Table 4.5: Centrally optimal vs. local optimal holdings in percentage

- The firm's centrally optimal trades are tilt toward U.S. and European bonds due to their lower volatility which result in a net bond position around 68%. However, after delegating the portfolio choices to un-coordinated agents, the new portfolio is now loaded much more heavily with equities, with the net bond positions reduced to a mere 37%. The overall efficiency loss in terms of risk-adjusted profit is a stunning -40%.

- Decentralized agents optimizing in isolation are only concerned about their own risk exposures. In particular, there is (i) no penalty for loading up on risk that the firm is already exposed to, and (ii) no incentive to hedge exposure that the firm might already have, due to the holdings of other agents.

- The firm's risk budget is a scarce commodity which has not been correctly priced.

In summary coordination is difficult due to a misalignment of incentives and lack of information, while the cost of not coordinating is substantial for the firm.

4.3.3 Swap Contract

A swap is a contract between two parties that specifies the size and direction of cash payments between them as a function of some observable factor (which is defined in the contract) [8]. Motivated by our example, we introduce swap contracts on the factors $W_1(t), \cdots, W_n(t)$ as a mechanism for transferring risks between agents in our system. Risk transfer changes the risk exposure of each of the agents and as a consequence their optimal investment behavior. With the introduction of contracts, it is natural to ask whether they can be chosen such that the resulting set of decentralized decisions is centrally optimal.

For illustration, consider Figure 4.1 where q_{ij} corresponds to the amount of cash transferred from Agent i to Agent j per unit change of the factor $W_j(t)$. For example, for a change $\Delta W_j(t)$, Agent 1 transfers $q_{12}\,\Delta W_j(t)$ to Agent 2 (and Agent 2 transfers $-q_{12}\,\Delta W_j(t)$ to Agent 1). In this case q_{12} defines a swap contract on the factor $W_j(t)$ between agents 1 and 2 and $(q_{12}+q_{13})\Delta W_j(t)$ is the net *payment* by Agent 1 from its swap positions with Agent 2 and 3. If $R_{ij}(t)$ denotes the net income received by Agent i per unit change in the factor $W_j(t)$ at time t, it follows that

$$\begin{aligned} R_{1j}(t)\Delta W_j(t) &= (-q_{12}-q_{13})\Delta W_j(t), \\ R_{2j}(t)\Delta W_j(t) &= (q_{12}-q_{23})\Delta W_j(t), \\ R_{3j}(t)\Delta W_j(t) &= (q_{13}+q_{23})\Delta W_j(t) \end{aligned}$$

where $\{R_{1j}(t), R_{2j}(t), R_{3j}(t)\}$ satisfies the property

$$R_{1j}(t) + R_{2j}(t) + R_{3j}(t) = 0.$$

since payments are internal cash transfers. More formally, we define an admissible system of cash transfers on factor $W_j(t)$ as follows:

Definition 4.3.1 (Admissible cash transfers)**.**
$R(t) \equiv \{R_{ij}(t), i = 1, \cdots, m,\ j = 1, \cdots, n\}$ *is admissible if* $R_{ij}(t)$ *is adapted to the BM*

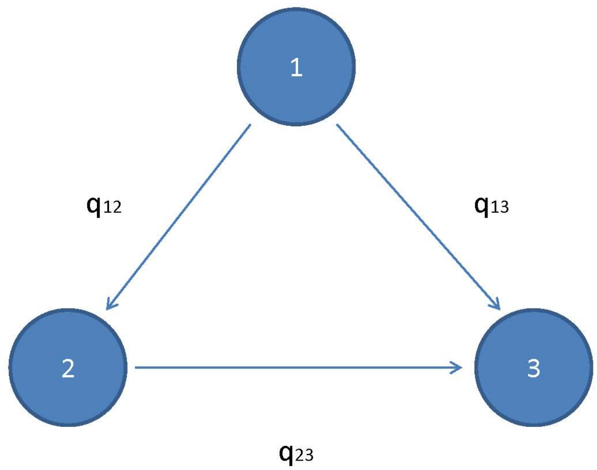

Figure 4.1: Swap contract between three agents

filtration $\{\mathcal{F}_t^W\}$ *and the budget balanced condition*

$$\sum_{i=1}^{m} R_{ij}(t) = 0,$$

is satisfied for all $t \in [0,\, T]$ *w.p. 1.*

Intuitively, Agent i knows $R_{ij}(t)$ $(j = 1, \cdots, n)$, and if there is a change $\Delta W_j(t)$ in factor j, he receives a net cash amount of $R_{ij}(t)\Delta W_j(t)$ from all other agents. Cash transfers is to modify Agent i's wealth, changing his instantaneous earnings to

$$\begin{aligned} \Delta x_i(t) &= \overbrace{\pi_i(t)\Big\{\alpha_i \Delta t + \sum_{j=1}^{n} \sigma_{ij}\Delta W_j(t) + \epsilon_i \Delta V_i(t)\Big\}}^{\text{risky asset holdings} = \pi_i(t)\frac{\Delta S_i(t)}{S_i(t)}} + \overbrace{\sum_{j=1}^{n} R_{ij}(t)\Delta W_j(t)}^{\text{swaps}}, \\ &= \pi_i(t)\alpha_i \Delta t + \sum_{j=1}^{n} \Big\{\pi_i(t)\sigma_{ij} + R_{ij}(t)\Big\}\Delta W_j(t) + \pi_i(t)\epsilon_i \Delta V_i(t). \end{aligned} \tag{4.18}$$

Clearly the swap redistributes exposures to risk factors among the agents. For example, if $\sigma_{ij} > 0$, then it becomes more expensive for Agent i to increase his holding in asset i if $R_{ij}(t) > 0$. Alternatively, if $R_{ij}(t) < 0$, then exposure to factor j has been transferred to other agents and there is additional incentive for Agent i to take on a larger position in asset i.

4.3.4 Agent i's Modified Problem

For a given transfer contract $\{R_{ij}(t)\}$, Agent i maximizes his expected terminal utility conditional on his own investment model and his assigned transfer payments $R_i(t) \triangleq \{R_{i1}(t), \cdots, R_{in}(t)\}$:

$$\begin{cases} \max_{\pi_i} \mathbb{E}_{\mathbb{P}}\Big[x_i(T)\Big] - \mathcal{R}_{(\theta_i,\widetilde{\theta}_i)}(x_i(T), R_{ij}(t)) \\ \equiv \max_{\pi_i} \min_{\gamma,\widetilde{\gamma}} \ \mathbb{E}_{\mathbb{Q}}\Big\{ x_i(T) + \sum_{j=1}^{n} \frac{\theta_{ij}}{2}\int_0^T |\gamma_{ij}(t)|^2 dt + \frac{\widetilde{\theta}_i}{2}\int_0^T |\widetilde{\gamma}_i(t)|^2 dt\Big\} \\ \text{subject to:} \\ dx_i(t) = \Big\{\pi_i(t)\Big[\alpha_i - \sum_{j=1}^{n}\gamma_{ij}(t)\sigma_{ij} - \widetilde{\gamma}_i(t)\epsilon_i\Big] - \sum_{j=1}^{n}\gamma_{ij}(t)\widetilde{R}_{ij}(t)\Big\}dt \\ \qquad\qquad + \sum_{j=1}^{n}\Big\{\pi_i(t)\sigma_{ij} + \widetilde{R}_{ij}(t)\Big\}d\widetilde{W}_j(t) + \pi_i(t)\epsilon_i d\widetilde{V}_i(t), \\ x_i(0) \text{ given.} \end{cases} \tag{4.19}$$

Though $R_i(t)$ is allowed to be any process adapted to the filtration generated by the factors $W_1(t), \cdots, W_m(t)$, in the special case when it is only a function of time, dynamic programming implies that the value function for Agent i is

$$\begin{cases} V_i(t, x_i) = x_i + H_i(t), \\ \text{where,} \\ 0 = \dot{H}_i(t) + \max_{\pi_i}\Big\{\pi_i\alpha_i - \frac{1}{2\widetilde{\theta}_i}(\pi_i\epsilon_i)^2 \\ \qquad - \sum_{j=1}^{n}\frac{1}{2\theta_{ij}}\Big[\pi_i\sigma_{ij} + R_{ij}(t)\Big]^2\Big\}, \\ H_i(T) = 0, \end{cases} \tag{4.20}$$

where the optimal portfolio $\pi_i(t|R_i(t))$ solves

$$\alpha_i - \frac{1}{\widetilde{\theta}_i}(\epsilon_i)^2\pi_i - \sum_{j=1}^{n}\frac{\sigma_{ij}}{\theta_{ij}}\Big(\pi_i\sigma_{ij} + R_{ij}(t)\Big) = 0. \tag{4.21}$$

4.3.5 Weak Duality

For every admissible set of contracts $R(t) \equiv [R_1(t) \cdots, R_m(t)]$, each agent formulates and solves his decentralized problem (4.19) for an optimal investment holding $\pi_i(\cdot \,|\, R_i(t))$, and has end of period wealth $x_i(T)$. Since contracts specify internal cash transfers, it is easy to show that the firm's net wealth is

$$V(R(t)) = x_1(T) + ... + x_m(T).$$

It is clearly of interest how the firm's net utility under decentralized control relates to the value function of (4.12) under optimal centralized control. The following result follows immediately from the observation that the decentralized investment policy $\{\pi_1(\cdot \mid R_1(t)), \cdots, \pi_m(\cdot \mid R_m(t))\}$ are admissible though not generally optimal for the centralized problem.

Proposition 4.3.1 (Weak Duality).
Let $R(t)$ be an arbitrary admissible swap contract, $\{\pi_1(\cdot \mid R_1(t)), \cdots, \pi_m(\cdot \mid R_m(t))\}$ the optimal decentralized investment policies under this contract, $V(R)$ the resulting expected utility for the firm (4.19), and V^* the value function for the centralized agent (4.12). Then

$$V(R(t)) \leq V^*. \tag{4.22}$$

Clearly, it is of interest to characterize the firm's optimal contract

$$\begin{cases} V(R^*(t)) \triangleq \max\limits_R \mathbb{E}_{\mathbb{P}}\Big[\sum_{i=1}^m x_i(T)\Big] - \mathcal{R}_{(\theta,\widetilde{\theta})}\Big(\sum_{i=1}^m x_i(T)\Big) \\ \text{subject to:} \\ \quad x_i(T) \equiv \text{optimal wealth of Agent } i \text{ under contract } R(t), \end{cases}$$

and to determine the efficiency loss relative to centralized optimality of the optimal contract.

4.4 Optimal Transfer Contracts

In this section, we construct an optimal contract using dynamic programming arguments and show that there is no efficiency loss relative to centralized optimality.

4.4.1 Optimal Contract: Conjecture

Let $\theta_{ij} > 0$ denote the risk tolerance of Agent i toward factor j, and $\theta_j = \theta_{1j} + \cdots + \theta_{mj}$ the risk tolerance for the firm toward fact j. Observing that

$$\begin{aligned}
& \pi(t)'\sigma \cdot [diag(\theta)]^{-1} \cdot \sigma'\pi(t) \\
= \;& \sum_{j=1}^n \frac{1}{\theta_j}\Big[\sum_{i=1}^m \pi_i(t)\sigma_{ij}\Big]^2 \\
= \;& \sum_{i=1}^m \sum_{j=1}^n \frac{1}{\theta_{ij}}\Big[\frac{\theta_{ij}}{\theta_j}\sum_{i=1}^m \pi_i(t)\sigma_{ij}\Big]^2 \\
= \;& \sum_{i=1}^m \sum_{j=1}^n \frac{1}{\theta_{ij}}\Big[\pi_i(t)\sigma_{ij} + \sum_{k\neq i}\Big(\frac{\theta_{ij}}{\theta_j}\pi_k(t)\sigma_{kj} - \frac{\theta_{kj}}{\theta_j}\pi_i(t)\sigma_{ij}\Big)\Big]^2,
\end{aligned}$$

the dynamic programming equations (4.14) for the centralized problem can be written as

$$
\begin{cases}
V(t,\,x) = x + H(t), \\
\text{where,} \\
0 = \dot{H}(t) + \max_{\pi} \sum_{i=1}^{m} \Big\{ \pi_i \alpha_i - \dfrac{1}{2\widetilde{\theta}_i} (\pi_i \epsilon_i)^2 \\
\qquad\qquad - \sum_{j=1}^{n} \dfrac{1}{2\theta_{ij}} \Big[\pi_i \sigma_{ij} + \sum_{k \neq i} \Big(\dfrac{\theta_{ij}}{\theta_j} \pi_k \sigma_{kj} - \dfrac{\theta_{kj}}{\theta_j} \pi_i \sigma_{ij} \Big) \Big]^2 \Big\}, \\
H(T) = 0.
\end{cases}
\tag{4.23}
$$

Let $\pi^*(t) = \{\pi_1^*(t), \cdots, \pi_m^*(t)\}$ denote the optimal centralized holdings in each of the m markets (the maximizer in (4.23)) and consider the modified system obtained by replacing some of the variables in (4.23) with their optimal quantities:

$$
\begin{cases}
V(t,\,x) = x + H(t), \\
\text{where,} \\
0 = \dot{H}(t) + \max_{\pi} \sum_{i=1}^{m} \Big\{ \pi_i \alpha_i - \dfrac{1}{2\widetilde{\theta}_i} (\pi_i \epsilon_i)^2 \\
\qquad\qquad - \sum_{j=1}^{n} \dfrac{1}{2\theta_{ij}} \Big[\pi_i \sigma_{ij} + \sum_{k \neq i} \Big(\dfrac{\theta_{ij}}{\theta_j} \pi_k^*(t) \sigma_{kj} - \dfrac{\theta_{kj}}{\theta_j} \pi_i^*(t) \sigma_{ij} \Big) \Big]^2 \Big\}, \\
H(T) = 0.
\end{cases}
\tag{4.24}
$$

By comparing the first order conditions, it is easy to show that the original system (4.23) and the modified system (4.24) are equivalent in that both have the same maximizer $\pi^*(t) = \{\pi_1^*(t), \cdots, \pi_m^*(t)\}$ and solution $H(t)$, and hence, both characterize the value function $V(t,\,x)$ and the optimal portfolio for the centralized problem (4.12).

The advantage of (4.24) is that it is separable; there is only one decision variable $\pi_i(t)$ in each term

$$
\pi_i \alpha_i - \frac{1}{2\widetilde{\theta}_i} (\pi_i \epsilon_i)^2 - \sum_{j=1}^{n} \frac{1}{2\theta_{ij}} \Big[\pi_i \sigma_{ij} + \sum_{k \neq i} \Big(\frac{\theta_{ij}}{\theta_j} \pi_k^*(t) \sigma_{kj} - \frac{\theta_{kj}}{\theta_j} \pi_i^*(t) \sigma_{ij} \Big) \Big]^2 .
$$

The leads us to defining $V_i(t,\,x_i)$ as the solution of

$$
\begin{cases}
V_i(t,\,x_i) = x_i + H_i(t), \\
0 = \dot{H}_i(t) + \max_{\pi_i} \Big\{ \pi_i \alpha_i - \dfrac{1}{2\widetilde{\theta}_i} (\pi_i \epsilon_i)^2 - \sum_{j=1}^{n} \dfrac{1}{2\theta_{ij}} \Big[\pi_i \sigma_{ij} + R_{ij}^*(t) \Big]^2 \Big\}, \\
H_i(T) = 0,
\end{cases}
\tag{4.25}
$$

where

$$
R_{ij}^*(t) = \sum_{k \neq i} \Big(\frac{\theta_{ij}}{\theta_j} \pi_k^*(t) \sigma_{kj} - \frac{\theta_{kj}}{\theta_j} \pi_i^*(t) \sigma_{ij} \Big), \tag{4.26}
$$

and allows us to write

$$V(t,x) = V_1(t,x_1) + ... + V_m(t,x_m),$$

for any $x_1, \cdots, x_m$ such that

$$\sum_{i=1}^{m} x_i = x.$$

The key observation, however, is that (4.25) is the dynamic programming equation for Agent i's problem (4.19) under transfer payments (4.26) and that the maximizer $\pi_i^*(t)$ in (4.25) equals the i^{th} component of the centralized agent's optimal allocation policy $\pi^* = \{\pi_1^*(t), \cdots, \pi_m^*(t)\}$, which suggests that with transfer contracts (4.26), decentralized agents optimally choose the centralized optimal policy.

4.4.2 Optimal Contract: Verification of Optimality

To establish optimality of (4.26), we need to show that it is admissible (as in Definition 4.3.1) and that the resulting collection of decentralized optimal policies coincides with the centralized optimal. Admissibility is straightforward, while optimality can be shown using the first order conditions.

Theorem 4.4.1. *Let $\theta_{ij} > 0$ denote the risk tolerance for Agent i toward factor j, $\theta_j = \theta_{1j} + \cdots + \theta_{mj}$ the risk tolerance for the firm toward factor j, and $\pi^*(t) = \{\pi_1^*(t), \cdots, \pi_m^*(t)\}$ defined by (4.15) denote the optimal policy for the centralized agent. Then the contract*

$$R_{ij}^*(t) = \sum_{k \neq i} \Big(\frac{\theta_{ij}}{\theta_j} \pi_k^*(t) \sigma_{kj} - \frac{\theta_{kj}}{\theta_j} \pi_i^*(t) \sigma_{ij} \Big) \tag{4.27}$$

is admissible and the optimal investment for Agent i equals the optimal investment $\pi^*(t)_i$ in market i by the centralized agent. The contract $R^*(t)$ is optimal and there is no efficiency loss relative to the optimal utility for the centralized agent

$$\max_R V(R) = V(R^*(t)) = V^*(t).$$

Proof. Admissibility of the contract (4.26) follows from the observation that is defines internal cash transfers between agents. To show that Agent i's optimal holding under (4.26) equals the component $\pi_i^*(t)$ of the optimal holding for the centralized agent, we need to show that $\pi_i^*(t)$ is the optimizer in the dynamic programming equation (4.25) for Agent i

$$\pi_i^*(t) = \arg\max_{\pi_i} G_i(\pi_i),$$

for each $i = 1, \cdots, m$, where

$$G_i(\pi_i) \triangleq \Big\{\pi_i\alpha_i - \frac{1}{2\widetilde{\theta}_i}(\pi_i\epsilon_i)^2 - \sum_{j=1}^{n}\frac{1}{2\theta_{ij}}\Big[\pi_i\sigma_{ij} + \sum_{k\neq i}\Big(\frac{\theta_{ij}}{\theta_j}\pi_k^*(t)\sigma_{kj} - \frac{\theta_{kj}}{\theta_j}\pi_i^*(t)\sigma_{ij}\Big)\Big]^2\Big\}.$$

Equivalently, we can show $G_i'(\pi_i^*(t)) = 0$. To see this, observe (using $\theta_j = \theta_{ij} + \sum_{k\neq i}\theta_{kj}$ and (4.26)) that

$$\begin{aligned} G_i'(\pi_i) &= \alpha_i - \frac{1}{\widetilde{\theta}_i}\pi_i\epsilon_i^2 - \sum_{j=1}^{n}\frac{\sigma_{ij}}{\theta_{ij}}\Big\{\pi_i\sigma_{ij} + \Big[\sum_{k\neq i}\frac{\theta_{ij}}{\theta_j}\pi_k^*(t)\sigma_{kj} - \pi_i^*(t)\sigma_{ij}\sum_{k\neq i}\frac{\theta_{kj}}{\theta_j}\Big]\Big\} \\ &= \alpha_i - \frac{1}{\widetilde{\theta}_{ij}}\pi_i\epsilon_i^2 - \sum_{j=1}^{n}\frac{\sigma_{ij}}{\theta_{ij}}\Big\{\pi_i\sigma_{ij} + \Big[\sum_{k\neq i}\frac{\theta_{ij}}{\theta_j}\pi_k^*(t)\sigma_{kj} - \pi_i^*(t)\sigma_{ij}\frac{1}{\theta_j}(\theta_j - \theta_{ij})\Big]\Big\} \\ &= \alpha_i - \frac{1}{\widetilde{\theta}_i}\pi_i\epsilon_i^2 - \sum_{j=1}^{n}\frac{\sigma_{ij}}{\theta_{ij}}\Big\{\pi_i\sigma_{ij} - \pi_i^*(t)\sigma_{ij} + \sum_{k=1}^{m}\frac{\theta_{ij}}{\theta_j}\pi_k^*(t)\sigma_{kj}\Big\}. \end{aligned}$$

It now follows that

$$G_i'(\pi_i^*(t)) = \alpha_i - \frac{1}{\widetilde{\theta}_{ij}}\pi_i(t)\epsilon_i^2 - \frac{1}{\theta_j}\sum_{j=1}^{n}\sigma_{ij}\sum_{k=1}^{m}\pi_k^*(t)\sigma_{kj} = 0,$$

where the last equality follows from the observation that $\pi^*(t) = [\pi_1^*(t), \cdots, \pi_m^*(t)]$ is optimal for the centralized agent and hence satisfies the first order conditions (4.16) for the centralized problem. We can therefore conclude that the collection of decentralized agents under the contract (4.26) make portfolio allocations that equal the optimal allocations the centralized agent. It follows that the firm's utility under (4.26) equals the optimal utility of the centralized agent, $V(R^*(t)) = V^*$. Weak duality (4.22) implies that (4.27) is the optimal contract and that there is no efficiency loss relative to the optimal utility for the centralized agent. □

One convenient property of the optimal contract (4.27) is that it is deterministic (a consequence of the firm's utility function in (4.12) being exponential), which implies that the optimal portfolio for each decentralized agent is characterized by (4.21).

Figure 4.2 illustrates the optimal transfer contract between three agents which in the case of Agent 1 given by

$$R_{1j}^*(t) = \underbrace{\Big\{\frac{\theta_{1j}}{\theta_j}v_{2j}^*(t) - \frac{\theta_{2j}}{\theta_j}v_{1j}^*(t)\Big\}}_{\text{swap with Agent 2}} + \underbrace{\Big\{\frac{\theta_{1j}}{\theta_j}v_{3j}^*(t) - \frac{\theta_{3j}}{\theta_j}v_{1j}^*(t)\Big\}}_{\text{swap with Agent 3}},$$

where $v_{ij}^*(t) = \pi_i^*(t)\sigma_{ij}$ denotes Agent i's exposure to factor $W_j(t)$. Agent 1's transfer R_{1j} is the sum of a payments associated with swap agreements with Agents 2 and 3. In each swap, both agents transfer a proportion of their exposure to factor j to the other agent, where proportions are determined by risk tolerances to that factor.

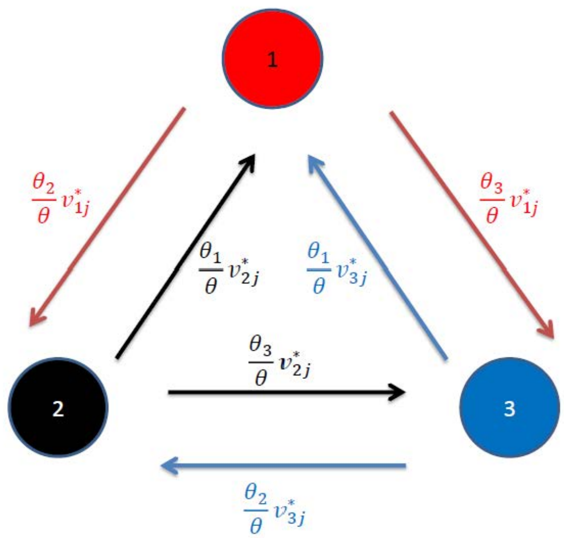

Figure 4.2: Optimal contracts: intuition

4.5 Decentralized Computation of Optimal Contracts

Theorem 4.4.1 characterizes the swap contracts under which decentralized agents optimally choose the centrally optimal holdings. We now turn to the question of computing these contracts. While one approach would be to solve the centralized problem directly, this is problematic because a centralized agent with the information required to solve the centralized problem typically does not exist (a key motivation for this paper), making it unsatisfying to call on such an agent to compute the optimal transfer contracts. In this section, we propose an iterative approach for computing optimal contracts that does not require a centralized agent and can be implemented without decentralized agents having to reveal private market information to others. Convergence to the optimal transfer contract will then be established.

The proposed algorithm makes the assumption that each agent is required to report his exposure to each risk factor to the bank's risk management system. For example, Agent i must report his exposure $v_{ij}(t) = \pi_i(t)\sigma_{ij}$ to each factor $W_j(t)$ (α_i and $\sigma_{i1}, \cdots, \sigma_{in}$ remain private). Observing that the optimal contract

$$R^*_{ij}(t) = \sum_{k \neq i} \Big(\frac{\theta_{ij}}{\theta_j} \pi^*_k(t)\sigma_{kj} - \frac{\theta_{kj}}{\theta_j} \pi^*_i(t)\sigma_{ij} \Big) = \sum_{k \neq i} \Big(\frac{\theta_{ij}}{\theta_j} v^*_{kj}(t) - \frac{\theta_{kj}}{\theta_j} v^*_{ij}(t) \Big)$$

depends on the optimal exposures $v^*_{ij}(t)$ of each agent to each factor, a natural idea is to iterate on the set of contracts and holdings.

Algorithm 4.5.1 is motivated by these considerations. In each iteration, decentralized agents solve their own optimization problems (4.19) conditional on some (suboptimal) transfer contract and report the resulting exposures to the firm. The firm updates the contracts and passes the relevant pieces to the agents, and the process repeats. It is important to note that a centralized agent with knowledge of the integrated market model is not required for implementation and that decentralized agents need not reveal private market information. Indeed, all each agent needs is the current contract $\widetilde{R}_i$ and his own market information, while the number of other agents, their models, contracts, risk-tolerances are irrelevant.

Algorithm 4.5.1.
Start with $R(t) \equiv 0$ and proceed one Agent at a time.

1. *Firm computes contract*

$$\widetilde{R}_{ij}(t) = \sum_{k\neq i} \Big\{ \frac{\theta_{ij}}{\theta_j} \widetilde{v}_{kj}(t) - \frac{\theta_{kj}}{\theta_j} \widetilde{v}_{ij}(t) \Big\}.$$

   using most recently reported exposures $\widetilde{v}(t) = \{\widetilde{v}_{ij}(t)\}$.

2. *If it is Agent i's turn to update, send him the contract*

$$\widetilde{R}_i(t) \equiv [\widetilde{R}_{i1}(t), \cdots, \widetilde{R}_{in}(t)].$$

3. *Agent i maximizes his objective*

$$\begin{cases} \max\limits_{\pi_i} \mathbb{E}_{\mathbb{P}}\Big[x_i(T)\Big] - \mathcal{R}_{(\theta_i,\widetilde{\theta}_i)}(x_i(T), \widetilde{R}_{ij}(t)) \\ \equiv \max\limits_{\pi_i} \min\limits_{\gamma,\widetilde{\gamma}} \ \mathbb{E}_{\mathbb{Q}}\Big\{ x_i(T) + \sum_{j=1}^{n} \frac{\theta_{ij}}{2} \int_0^T |\gamma_{ij}(t)|^2 dt + \frac{\widetilde{\theta}_i}{2} \int_0^T |\widetilde{\gamma}_i(t)|^2 dt \Big\} \\ \textit{subject to:} \\ dx_i(t) = \Big\{ \pi_i(t) \Big[\alpha_i - \sum_{j=1}^{n} \gamma_{ij}(t)\sigma_{ij} - \widetilde{\gamma}_i(t)\epsilon_i \Big] - \sum_{j=1}^{n} \gamma_{ij}(t) \widetilde{R}_{ij}(t) \Big\} dt \\ \qquad + \sum_{j=1}^{n} \Big\{ \pi_i(t)\sigma_{ij} + \widetilde{R}_{ij}(t) \Big\} d\widetilde{W}_j(t) + \pi_i(t)\epsilon_i d\widetilde{V}_i(t), \\ x_i(0) \textit{ given.} \end{cases}$$

   and reports his new exposure $\widetilde{v}^*_{ij}(t) = \widetilde{\pi}^*_i(t)\sigma_{ij}$ to the firm.

4. *Move on to the next Agent and repeat the process.*

Observe that if $\widetilde{R}_{ij}(t)$ in step 2 is deterministic, then the optimal portfolio for Agent i in step 3 is also deterministic (and characterized by (4.21)), which leads to a deterministic transfer contract $\widetilde{R}_{ij}(t)$ in the subsequent iteration. That is, if we start with a deterministic contract, they remain deterministic (though time varying) throughout the algorithm. Finally, observe that the dimension of Agent i's problem is lower than that of the centralized agent.

4.5.1 Convergence

We now present results that guarantee convergence of the Algorithm 4.5.1 to the optimal set of contracts.

Theorem 4.5.1. *Let* $\{\pi_i^k(t)\}$ *be the sequence of optimal holdings computed by Algorithm 4.5.1. The* $\{\pi_i^k(t)\}$ *converges strong to* $\pi_i^*(t)$*, namely,*

$$\lim_{k\to\infty} \|\pi_i^k(t) - \pi_i^*(t)\| = 0, \quad \forall i = 1, ..., m.$$

Correspondingly,

$$\lim_{k\to\infty} \|V(R^k(t)) - V^*\| = 0.$$

Proof. Given the contract

$$\begin{aligned} \widetilde{R}_{ij}(t) &= \sum_{k\neq i} \Big\{ \frac{\theta_{ij}}{\theta_j} \widetilde{v}_{kj}(t) - \frac{\theta_{kj}}{\theta_j} \widetilde{v}_{ij}(t) \Big\}, \\ &= \sum_{k\neq i} \Big\{ \frac{\theta_{ij}}{\theta_j} \widetilde{\pi}_k(t)\sigma_{kj} - \frac{\theta_{kj}}{\theta_j} \widetilde{\pi}_i(t)\sigma_{ij} \Big\}, \end{aligned}$$

Agent i updates his holding from $\widetilde{\pi}_i(t)$ to

$$\Rightarrow \widetilde{\pi}_i^*(t) = \frac{\alpha_i - \sum_{j=1}^n \frac{1}{\theta_{ij}} \sigma_{ij} \widetilde{R}_{ij}(t)}{\frac{1}{\widetilde{\theta}_i}\epsilon_i^2 + \sum_{j=1}^n \frac{1}{\theta_{ij}} \sigma_{ij}^2},$$

It can be shown that

$$\widetilde{\pi}_i^*(t) = \Big(1 - \lambda\Big)\widetilde{\pi}_i(t) + \lambda \left[\frac{\alpha_i - \sum_{k\neq i} \widetilde{\pi}_k(t) \sum_{j=1}^n \frac{1}{\theta_j} \sigma_{kj}\sigma_{ij}}{\frac{1}{\widetilde{\theta}_i}\epsilon_i^2 + \sum_{j=1}^n \frac{1}{\theta_{ij}} \sigma_{ij}^2} \right],$$

where,

$$\lambda = \frac{\frac{1}{\widetilde{\theta}_i}\epsilon_i^2 + \sum_{j=1}^n \frac{1}{\theta_{ij}} \frac{\theta_{ij}}{\theta_j} \sigma_{ij}^2}{\frac{1}{\widetilde{\theta}_i}\epsilon_i^2 + \sum_{j=1}^n \frac{1}{\theta_{ij}} \sigma_{ij}^2},$$

which is a convex combination of the original holding $\widetilde{\pi}_i(t)$ and the holding obtained by optimizing the *centralized* objective function over $\pi_i(t)$ with other holdings kept fixed at $\widetilde{\pi}_j(t)$. The key observation is that each agent's update is equivalent to a so-called successive over-relaxation step of the firm's objective function [7], which gives us convergence. □

Returning to our running example, Figure 4.3 plots the convergence of Algorithm 4.5.1. While a large number of iterations is needed to close the gap from -80% to 0%, the substantial improvement observed after a small number of rounds is promising. Finally Table 4.6 shows the optimal contract.

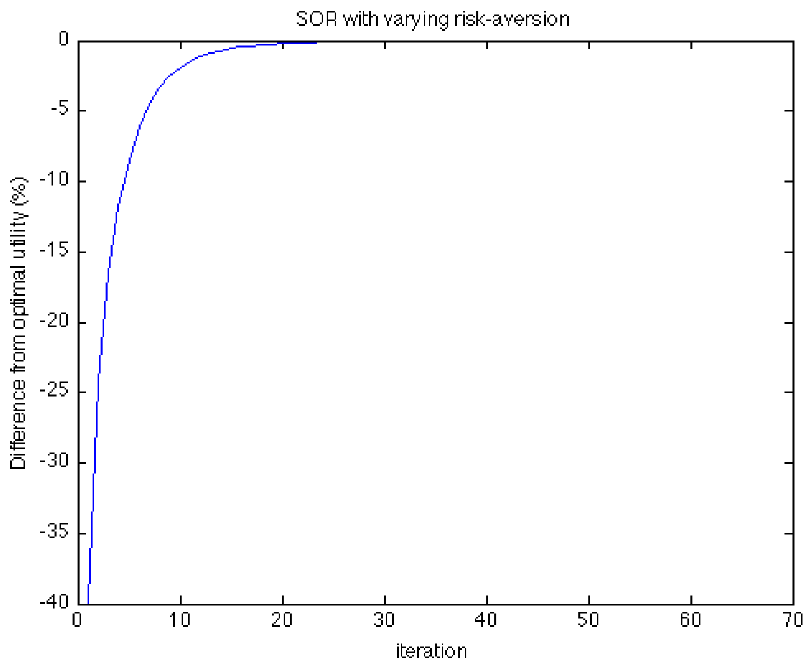

Figure 4.3: Deviation from centrally optimal utility converges from -40% to 0%.

	R^*							
1	-0.41	0.76	0.59	0.63	0.37	0.43	0.91	0.86
2	0.28	-1.26	-0.11	0.12	0.06	0.30	0.59	0.56
3	0.36	0.26	-0.34	0.27	0.23	0.36	0.45	0.44
4	0.21	0.09	0.07	-1.34	0.04	0.27	0.36	0.35
5	0.11	0.17	0.10	0.13	-0.36	0.15	0.30	0.29
6	-0.27	0.01	0.04	0.13	-0.14	-0.92	0.16	0.16
7	0.04	0.11	-0.06	0.10	0.02	-0.07	-1.12	-0.49
8	-0.33	-0.15	-0.29	-0.04	-0.21	-0.50	-1.65	-2.17

Table 4.6: Contracts for each of the 8 agents

4.6 Summary

In this chapter we study the problem of decentralized dynamic portfolio choice as a decentralized stochastic optimal control problem. We show that centralized efficiency can be achieved by introducing swap contracts between agents and provide an explicit characterization of the optimal contract. We also present an iterative algorithm for computing optimal swap contracts that it can be implemented without any of the agents having to reveal private information about their models. Our approach is not only novel, but also it offers a significant advantage over the benchmarking approach more conventionally employed in the finance literature. The benchmarking approach makes the unrealistic assumption that an all knowing centralized agent is available to construct the optimal performance benchmarks, which is entirely unnecessary in our approach and what we consider a major motivation for requiring decentralized control in the first place.

Chapter 5

Risk Attribution, Sensitivity Analysis, and Fair Allocation

5.1 Introduction

In this chapter we seek to understand the following important issues concerning efficient implementation and management of the swap contracts we introduce in the previous chapter,

- *Risk attribution*: we use factor models in the previous chapter to describe a firm's risk exposure, as it is commonly done in practice. In real-life applications, the number of risk factors can be quite large. Therefore, implementing and managing a large number of swap contracts efficiently can become a challenging issue. In Section 5.2 we discuss how to utilize risk attribution technique to decompose a firm's total risk exposure into different risk factors, and how to utilize such information to differentiate and prioritize the implementation of the swap contracts among different risk factors.

- *Sensitivity analysis*: from the case study used in the previous chapter, we see that the optimal swap contracts can help recover efficiency loss as much as 40%. However, this gain could be sensitive to certain underlying assumptions, such as the degree of correlation among the different markets. Moreover, model uncertainty is a prevailing issue in any quantitative modeling work that can undermine the effectiveness of our proposed solution. In Section 5.3 we analyze the sensitivity or robustness of the swap contracts' efficiency gain to the underlying assumptions as well as the issue of model uncertainty.

- *Fair allocation*: the main objective of the swap contracts is to align incentives between agents and the central manager. However, imposing swap contracts does not automatically guarantee participation. The overall utility gain needs to be redistributed among the agents in a fair manner and to provide sufficient incentives to the agents to secure their participation, in the sense that they will take on the swap contracts willingly. In Section 5.4 we discuss how to extend the swap contracts to include a participation bonus (or fee) to achieve fairness in allocation and to encourage agent participation.

5.2 Risk Attribution

In real-world applications the number of risk factors can be quite large, leading to a large number of swap contracts, which could be cumbersome to manage. A practical question would be, is it possible to contract only on a small number of risk factors, but gain most of the benefits? The key lies in the ability to carry out proper risk attribution. Recall in Section 4.1 that we model a firm's risk exposure using independent risk factors (plus idiosyncratic risk), which can be constructed via principal component analysis. More specifically, for the single period case, the firm's common risk exposure can be easily decomposed

$$\frac{1}{2}\pi(t)'\sigma \cdot [diag(\theta)]^{-1} \cdot \sigma'\pi(t) = \frac{1}{2}\sum_{j=1}^{n}\frac{1}{\theta_j}[\sum_{i=1}^{m}\sigma_{ij}\pi_i(t)]^2, \tag{5.1}$$

where the total common risk exposure can be expressed as a sum of the risk contribution from each independent risk factors. This provides us with a mean to attribute the total common risk exposure into the n risk factors. Naturally, after attribution we can rank the risk factors according to their contributions to the total common risk and prioritize the implementation of swap contracts accordingly.

Table 5.1 shows that factor 7 and 8 are the top two risk factors, each contributing above 20% of the total common risk exposure. Consequently, if we write only a single contract on either factor 7 or 8, the amount of utility gain is close to 18%, therefore a large chunk of the maximum possible gain of 40% can be achieved by writing a single contract. Furthermore, we can see the ranking of the risk factors according to their risk attribution is very consistent with the ranking of the utility gain by writing a single contract on each of them.

The fact that we can easily attribute risk into different factors allows us to rank the relative importance of the risk factors. When the number of risk factors is large, this offers great flexibility to investors to selectively construct swap contracts to gain most of the benefits without a full-blown implementation. In other words, investors have the flexibility to trade off optimality for operational efficiency and to reduce complexity in managing the swap positions.

5.3 Sensitivity Analysis

In the benchmark case (Table 4.3), we assume the common risk factors account for 70% of the total risk. It is this part of the risk that requires coordination among the agents to avoid over-concentration of risk, and we show that potential efficiency loss without coordination can be as high as -40%. It is not surprised that the extent of efficiency loss is sensitive to the assumption of the proportion of risk that is due to the common risk factors. When the proportion of common risk is high, the agents' markets are more dependent (or correlated), consequently coordination becomes more critical, and vice versa.

Table 5.2 runs a simple sensitivity analysis of the amount of efficiency loss (4^{th} column) against the proportion of common risk (1^{st} column). When the common risk is at a high 90%, the efficiency loss can reach as high as -48.4%. On the other hand, when the common risk is at a low 10% (i.e. the different markets have little common exposure), the efficiency loss shrinks to a mere -6.2% (though this in financial terms is still a huge amount). Given the rather wide range of possible efficiency loss, a cautious reader may question the importance of introducing swap contracts. We claim that it is still crucial that a firm with decentralized management structure adopt our proposed coordination mechanism, and we offer two supporting arguments:

- *Average case*: using the simplest term, we see that on average the efficiency loss is around -30%, considering the proportion of common risk in the range of 10% to 90%. This by all means is a enormous amount, and in reality it is not untypical to see very high correlations among seemingly different markets or strategies, such as the infamous contagion effects evidenced in several recent financial crisis (e.g. sub-prime meltdown, tech bubble, LTCM, and etc).
- *Model uncertainty*: the second argument is slightly more subtle. We see that in Table 5.2 the really optimistic case with a 10% common risk has a relatively small efficiency loss of -6.2%. However, given the complexity and interconnectedness of the different markets, can one be so sure that the common exposure is only a mere 10%? In any quantitative work, model uncertainty (or estimation error) is far too common, and is often one of the most critical aspects affecting the usefulness of any model in assisting decision making. Here by model uncertainty we specifically refer to the uncertainty concerning the proportion of the common risk, which can range from 0% to 100%. There are two simple ways of evaluating the impact of model uncertainty. First is the maximum loss of assuming the wrong model, which is defined as the maximum possible loss when the investor picks one model (e.g. a very low 10% common exposure), but nature turns out to the worst case (e.g. a very high 90% common exposure). Second is the mean loss of assuming the wrong model, which is defined as the average loss when the investor picks one model, and the average loss

Factor	Risk attribution	Single contract gain
8	23.6%	17.6%
7	22.7%	17.9%
6	15.0%	7.0%
5	10.1%	2.1%
2	9.2%	4.2%
4	8.9%	5.7%
3	6.3%	2.0%
1	4.3%	2.0%

Table 5.1: Risk attribution and gain from writing a single contract: risks are attributed into orthogonal factors via principal component analysis, and contract gain from single contract has consistent ranking with the risk attribution, i.e. the factor that contribute more risk will tend to provide more utility gain by writing contract on it.

when the nature varies across all possible models.

Take a look at Figure 5.1, the red curve (right axis) shows that the maximum or mean loss has a bell-curve shape, which means that at the two extremes the user in general suffers heavier loss than assuming moderate cases. Therefore in the presence of model uncertainty, it is in general safer to assume a moderate level of common risk, which can protect investor from massive loss that can go as much as -44% simply by assuming the wrong model. Correspondingly the blue curve (left axis) shows that by assuming moderate level of common exposure in the region of 40% to 70%, the amount of possible efficiency loss is in the region of -25% to -40%. This shows that in the presence of model uncertainty (which is far too common), the amount of potential efficiency loss remain elevated. It is therefore a more pressing issue to introduce swap contracts. Similarly, Figure 5.2 plots the entire surface of the possible losses due to model uncertainty, without taking the max or mean as in Figure 5.1. This allows us to examine the curvature of the loss due to model uncertainty. Again it is clear that the curvature is steeper at both extremes and much smoother in the central area, which implies that the results will tend to more robust to model uncertainty when we assume moderate level of common risk.

Common Risk (%)	Central Opt.	Dentral Opt.	Efficiency loss (%)
90%	8.73	4.50	-48.4%
80%	8.58	4.82	-43.8%
70%	8.64	5.20	-39.9%
60%	8.79	5.63	-36.0%
50%	9.01	6.16	-31.7%
40%	9.29	6.79	-26.9%
30%	9.62	7.58	-21.2%
20%	10.02	8.58	-14.4%
10%	10.48	9.82	-6.2%

Table 5.2: Sensitivity analysis of the efficient loss with respect to the proportion of common risk. We vary the proportion of common exposure from a high 90% to low 10%, and the correspond efficiency loss changes from -48.4% to -6.2%.

5.4 Fair Allocation

The final issue we would like to address is fairness in allocation and how to ensure enough incentives are provided to agents to ensure participation. The swap contracts we introduce address the fundamental issue of restoring efficiency for the overall system. However, we do not address utility gain or loss at the individual agent level. In order to achieve maximum risk-adjusted returns for the firm, individual agents will have to bear different risk and return profiles according to the overall system characteristics (e.g. correlation structure).

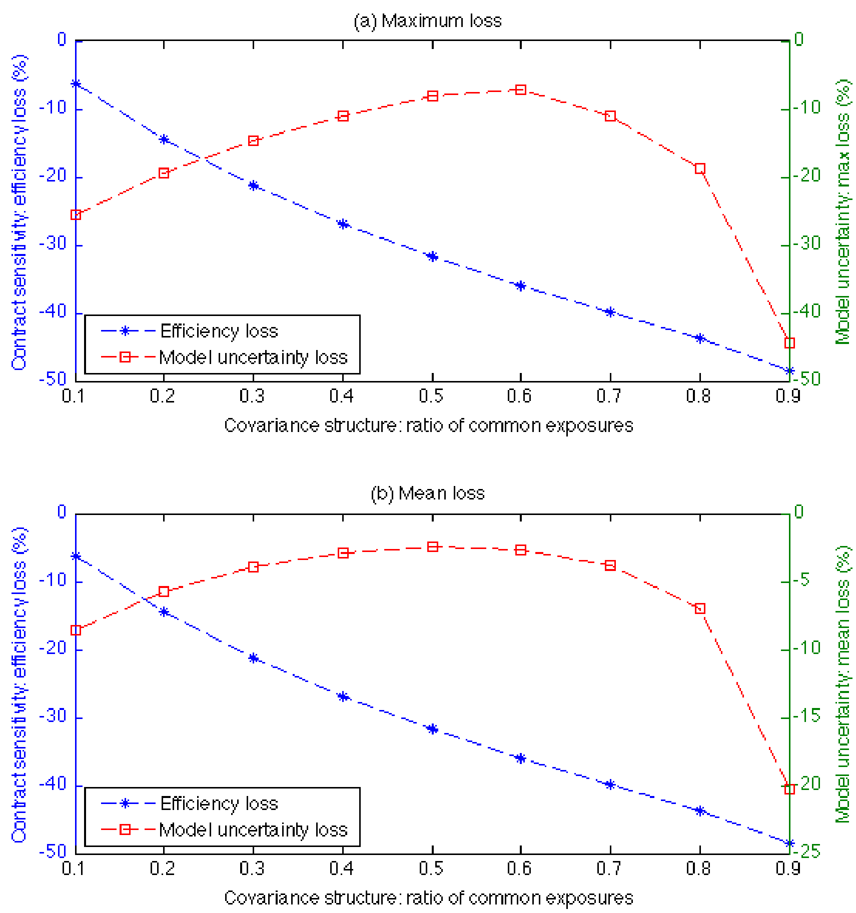

Figure 5.1: Sensitivity analysis in the presence of model uncertainty: blue curve (left axis) plots the efficiency loss, and red curve (right axis) plots the loss due to model uncertainty, against the x-axis where we vary the ratio of common exposure. Top panel shows the maximum loss (i.e. worst case), while the bottom panel shows the mean loss (i.e. average case).

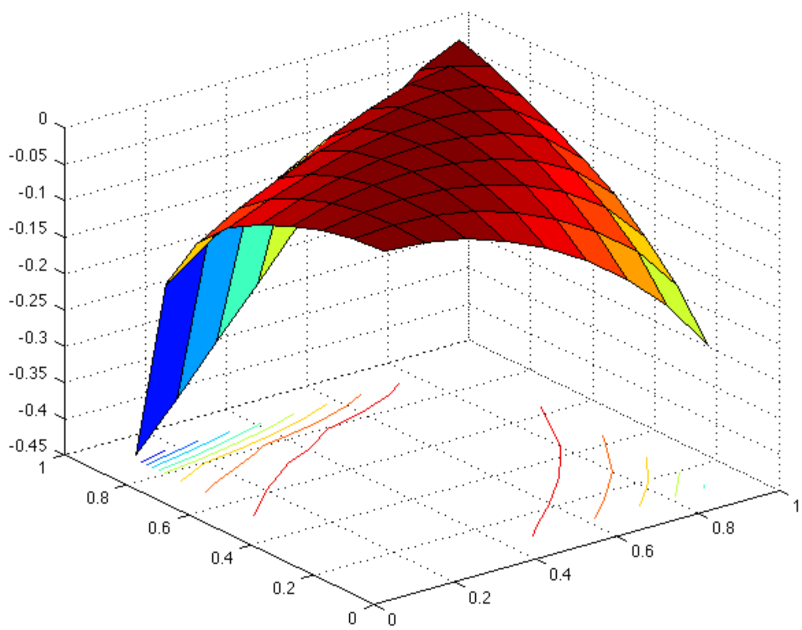

Figure 5.2: Sensitivity analysis in the presence of model uncertainty (curvature): this plots the entire surface of loss due to model uncertainty, without taking max or mean as in Figure 5.1.

Said differently, the added swap positions can cause either gain or loss at individual agent level. Therefore, the overall gain in efficiency will have to re-allocated to individual agents in a way that satisfy individual rationality constraints. Furthermore, the re-allocation need to be done in a fair manner.

Re-allocation of efficiency gain can be done by a simple extension of the swap contracts. Specifically, we can introduce a participation bonus (or fee) to be added to the swap contract. Denote this term as $\delta_i(t)$, we require $\sum_i \delta_i(t) = 0$, which implies that participation bonus (or fee) is simply a re-allocation of the utility gain. Agent i's change in wealth now becomes

$$\begin{aligned}
\Delta x_i(t) &= \overbrace{\pi_i(t)\Big\{\alpha_i \Delta t + \sum_{j=1}^{n} \sigma_{ij}\Delta W_j(t) + \epsilon_i \Delta V_i(t)\Big\}}^{\text{risky asset holdings } = \pi_i(t)\frac{\Delta S_i(t)}{S_i(t)}} + \overbrace{\sum_{j=1}^{n} R_{ij}(t)\Delta W_j(t)}^{\text{swaps}} + \overbrace{\delta_i(t)}^{\text{participation bonus (or fee)}}, \\
&= \pi_i(t)\alpha_i \Delta t + \sum_{j=1}^{n}\Big\{\pi_i(t)\sigma_{ij} + R_{ij}(t)\Big\}\Delta W_j(t) + \pi_i(t)\epsilon_i \Delta V_i(t) + \delta_i(t), \qquad (5.2)
\end{aligned}$$

where the participation bonus (or fee) $\delta_i(t)$ does not affect the optimal decisions, but act to provide incentives to the agents to encourage participation.

Table 5.3 shows that without coordination, the system can only achieve risk-adjusted return of 5.20 (2^{nd} column), with optimal swap contract, the risk-adjusted return can be improved to 8.64 (3^{rd} column). However, one can see that without fair re-allocation it can be net gain or loss at the individual agent level. A simple and fair allocation is to split the system gain equally among the agents, which leads to on average 39.9% gain over all agents (5^{th} column), with $\delta_i(t)$ the participation bonus (or fee) shown in last column. By adding this simple participation bonus (or fee), we ensure that we provide sufficient incentive to each individual agents to secure their participation.

5.5 Summary

In this chapter we discuss several important issues concerning efficient implementation and management of the swap contracts. First, we show that our choice of using independent risk factors via principal component analysis allows us to easily attribute the common risk exposure into individual risk factors. Consequently we can prioritize implementation of the swap contracts among different risk factors according to their contributions to the total risk. Risk attribution is particularly useful when there is a large number of common risk factors, and we show that writing partial contract on important risk factors can achieve a significant portion of the efficiency gain. This allows investor to easily trade off optimality for operational efficiency. Second, we discuss the sensitivity of contract gain to underlying assumptions, in particular the assumption of the proportion of common risk. We show that thought the contract gain can be rather sensitive to the underlying assumption, on average the potential gain can be as large as 30%. Furthermore, in the presence of model uncertainty, it is safer to assume a moderate level of common risk exposure, which provide further support of our claim that swap contracts can achieve substantial efficiency gain in the region of 25% to 40%. Lastly, we discuss the fairness issue. In order to encourage agents to take on swap contracts and participate in our proposed coordination framework, sufficient incentives need to be provided. We show that we can easily extend our swap contracts by adding a participation bonus (or fee), which implements an equal division on the overall system efficiency gain among the agents. This is a simple and fair re-allocation of the final wealth, and clearly provides sufficient incentives to secure agents' participation.

Agent	Local Opt.	Central Opt. (no re-allocation)	Central Opt. (fair re-allocation)	Improvement	δ_i
1	1.01	-0.75	1.44	30.0%	2.19
2	0.87	0.49	1.30	33.0%	0.81
3	0.34	-0.33	0.77	55.7%	1.10
4	0.33	0.58	0.76	56.4%	0.19
5	0.32	0.04	0.75	57.6%	0.71
6	0.50	1.02	0.93	46.2%	-0.09
7	0.70	2.03	1.13	38.1%	-0.90
8	1.12	5.56	1.55	27.8%	-4.01
Sum	5.20	8.64	8.64	39.9% (overall)	0.00

Table 5.3: Local optimal risk-adjusted profit versus central optimal risk-adjusted profit with and without fair re-allocation: note that the participation bonus (or fee) $\delta_i(t)$ adds up to zero, also the fair re-allocation ensures all agents receive equal share of the efficiency gain.

Chapter 6

Conclusion and Future Work

In this dissertation we study the fundamental issues of how to efficiently manage large-scale and multi-agent stochastic dynamic systems. In particular, we study how to design coordination mechanisms that would optimize system performance, and mechanism that can be implemented in a practical manner that respect agent level private information.

We study two classes of closely related problems. In the first part of this dissertation we study decentralized control of a general class of stochastic dynamic resource allocation problems. We model a stochastic system that is managed by multiple decentralized agents, who allocate system resource to satisfy customer demands that arrive stochastically over time. We introduce the notion of revenue transfer contract, which generalizes the notion of dual prices to stochastic and dynamic setting, and show that there exists optimal revenue transfer contracts that would properly coordinate agents' control policies to achieve central optimality. We show that the optimality is robust to potential model mis-specification in the sense that each agent typically does not know the modeling details of other agents but need to account for their actions in his own model. We further provide an iterative algorithm that can be used to construct the optimal revenue transfer contracts in decentralized manner, which does not require agents to reveal sensitive private information. A nice and somewhat surprising result is that optimality and convergence of the iterative algorithm does not require the typical assumption of convexity.

In the second part of the dissertation, we study a class of related problem but specializes to portfolio and risk management. We model a financial institution that trades in multiple correlated markets, and have to employ multiple agents to manage investment decision in the different markets due to the need for specialization. We show that un-coordination of agents can lead to substantial loss because of the loss of diversification and over-concentration in certain common risk factors. We design a rather unique coordination mechanism, which we coin as swap contract. Swap contract conceptually is similar to benchmarking that is often employed by investment firms, but differs substantially in that it can be constructed in decentralized manner without requiring an all-knowing central agent. Swap contract introduces a internal system of revenue transfer (or return sharing) that facilitates internal risk sharing. We show that optimal swap contract exists and can be constructed again in decentralized manner to guide individual agents to make investment decisions that would be optimal for the overall firm. These results offer new insights to decentralized investment management and a significant extension to the benchmarking literature in Finance.

We now briefly discuss several important future research directions:

- *Approximation methods*: we devote the entire chapter 3 discussing approximation methods for decentralized resource allocations. Computational issues concerning optimal control of stochastic dynamic system remain a significant challenge. This is further complicated by the need of decentralized control, where often iterative algorithms/schemes need to be employed. We discuss several approximations methods in chapter 3 that are adapted from their centralized counter-part. An interesting and important future work will be to apply these methods to large-scale systems and conduct empirical studies.

- *Decentralized learning*: the models in this dissertation assume that every agent know their local parameters (e.g. future demands) perfectly well. An important extension is to allow agents to learn their model parameters using realized data in a decentralized way. It will be interesting to see whether coordination can still be achieved in this setting and the impact of learning on the coordination mechanism.

- *Incentive issues*: in this dissertation we assume agents are price-takers or competitive, an interesting and important extension is to study different strategic behaviors that would allow agents to manipulate their reported information. This will extend our study into the dynamic game-theory territory.

Bibliography

[1] D. Adelman. Dynamic bid-prices in revenue management. *Operations Research*, 55(4):647–661, 2007.

[2] D. Adelman and A.J. Mersereau. Relaxations of weakly coupled stochastic dynamic programs. *Operations Research*, 56(3):712–727, 2008.

[3] K. J. Arrow and L. Hurwicz, editors. *Studies in Resource Allocation.* Cambridge University Press, Cambridge, UK, 1977.

[4] P. Artzner, F. Delbaen, J.M. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9(203-228), 1999.

[5] P. Barrieu and N. El Karoui. Optimal risk transfer. *Finance*, 25(31-47), 2004.

[6] D.P. Bersekas. Separable dynamic programming and approximate decomposition methods. *IEEE Transactions on Automatic Control*, 52(2):911–916, 2007.

[7] D.P. Bertsekas and J.N. Tsitsiklis. *Parallel and Distributed Computation: Numerical Methods.* Athena Scientific, Belmont, Massachussets, 1997.

[8] T. Bjork. *Arbitrage Theory in Continuous Time.* Oxford Univeristy Press, 2009.

[9] A. Boyd. Airline alliance. *OR/MS Today*, 1998.

[10] T. Brennan, A. Lo, and T. Nguyen. Portfolio theory: A review. MIT LFE Working Paper No. 2011–04–123., 2010.

[11] G. Cachon. *Supply chain coordination with contracts.* Handbooks in Operations Research and Lanagement Science: Supply Chain Management. 1998.

[12] H. Cai and A.E.B. Lim. Decentralized control of a multi-agent stochastic dynamic resource allocation problems in discrete time. *Proceedings of the* 50^{th} *IEEE Conference on Decision and Control*, pages 6400–6406, 2011.

[13] H. Cai and A.E.B. Lim. Decentralized control of stochastic dynamic resource allocation problems. Working Paper, IEOR Department, University of California (Berkeley), 2011.

[14] W.L. Cooper. Asymptotic behavior of an allocation policy for revenue management. *Operations Research*, (50):720–727, 2002.

[15] E.J. Elton and M.J. Gruber. Optimum centralized portfolio construction with decentralized portfolio management. *Journal of Financial and Quantitative Analysis*, 39(3):481–494, 2004.

[16] H. Föllmer and A. Schied. *Stochastic Finance: An Introduction to Discrete Time.* De Gruyter Studies in Mathematics, 2002.

[17] P. Grover, S.Y. Park, and A. Sahai. The finite-dimensional witsenhausen counterexample. *IEEE Transactions on Automatic Control*, Forthcoming.

[18] Y.C. Ho. Team decision theory and information structures. *Proceedings of the IEEE*, 68(6):644–654, 1980.

[19] Y.C. Ho. Review of the witsenhausen problem. *Proc. 47th IEEE Conference on Decision and Control*, pages 1611–1613, 2009.

[20] X. Hu, R. Caldentey, and G. Vulcano. Revenue sharing in airline alliances. *Management Science*, Forthcoming, 2012.

[21] L. Hurwicz K. J. Arrow and H. Uzawa, editors. *Studies in Linear and Nonlinear Programing.* Stanford University Press, 1958.

[22] B. Kurtaran. Dynamic two person two objective control problems with delayed sharing information pattern. *IEEE Transactions on Automatic Control*, 22:659–661, 1977.

[23] P. Li, A.E.B. Lim, and J.G. Shanthikumar. Optimal risk transfer for traders with germs. *Insurance: Mathematics and Economics*, 47(1):1–12, 2010.

[24] P. Li, A.E.B. Lim, and J.G. Shanthikumar. Decentralized control of a stochastic multiagent queueing system. *IEEE Transactions on Automatic Control*, Forthcoming, 2012.

[25] J. Marschak. Elements for a theory of teams. *Management Science*, 1:127–137, 1955.

[26] R.C. Merton. Lifetime portfolio selection under uncertainty: the continuous time case. *Rev. Econ. Stat.*, 51:247–257, 1969.

[27] C.C. Moallemi and B. Van Roy. Resource allocation via message passing. *INFORMS Journal of Computing*, 2010.

[28] R. Radner. Team decision problems. *Annals of Mathematical Statistics*, 33(3):857–881, 1962.

[29] A. Rantzer. Dynamic dual decomposition for distributed control. Working paper, Lund University, 2009.

[30] M. Rotkowitz and S. Lall. A characterization of convex problems in decentralized control. *IEEE Transactions on Automatic Control*, 51(2):274–286, 2006.

[31] N.R. Sandell and M. Athans. Solution of some nonclassical lqg stochastic decision problems. *IEEE Transactions on Automatic Control*, 19:108–116, 1974.

[32] N.R. Sandell, P. Varaiya, M. Athans, and M.G. Safonov. Survey of decentralized control methods for large scale systems. *IEEE Transactions on Automatic Control*, 23:108–128, 1978.

[33] W.F. Sharpe. Decentralized investment management. *Journal of Finance*, 36(2):217–234, 1981.

[34] D. Simchi-Levi, D.D. Wu, and Z.-J. Shen. *Supply chain analysis in the eBusiness era*. Kluwer, 2004.

[35] R.W. Simpson. Using network flow techniques to find shadow prices for market and seat inventory control. Technical report, Massachusetts Institute of Technology, 1989.

[36] K. Talluri and G. van Ryzin. *The theory and practice of revenue management*. Kluwer Academic Publishers, 2004.

[37] J.H. van Binsbergen, M.W. Brandt, and R.S.J. Koijen. Optimal decentralized investment management. *Journal of Finance*, 63(4):1849–1895, 2008.

[38] J.H. van Binsbergen, M.W. Brandt, and R.S.J. Koijen. Decentralized decision making in investment management. The Oxford Handbook of Quantitative Asset Management, 2010.

[39] P. Varaiya and J. Walrand. On delayed sharing patterns. *IEEE Transactions on Automatic Control*, 23:388–394, 1978.

[40] B. Vinod. Alliance revenue management. *Journal of Revenue and Pricing Management*, (4):66–82, 2005.

[41] E.L. Williamson. *Airline network seat control*. PhD thesis, Massachusetts Institute of Technology, 1992.

[42] H.S. Witsenhausen. A counterexample in stochastic optimum control. *SIAM Journal on Control*, 6(1):138–147, 1968.

[43] H.S. Witsenhausen. Separation of estimation and control for discrete time systems. *Proc. IEEE*, (59):1557–1566, 1971.

[44] C.P. Wright, H. Groenevelt, and R.A. Shumsky. Dynamic revenue management in airline alliances. *Transportation Science*, 44(1):15–37, 2010.