Title
New Operations Research Models for Emerging Problems in Service Operations

Author
Hassanzadeh Kalshani, Ali

Publication Date
2021

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,
IRVINE

New Operations Research Models for Emerging Problems in Service Operations

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Management

by

Ali Hassanzadeh Kalshani

Dissertation Committee:
Associate Professor Luyi Gui, Chair
Associate Professor John Turner, Chair
Professor Shuya Yin

2021

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ALGORITHMS

ACKNOWLEDGMENTS

Earning a doctoral degree and diving deep into some of the challenging topics of my field has always been my goal as a student, which seemed unattainable at times. I am excited that my Ph.D. journey has come to an end after successfully defending my dissertation and I am extremely grateful to all those who helped me, supported me, and enlightened me along the way with their guidance and wisdom. First, I would like to extend my utmost appreciation towards my advisors, professors John Turner and Luyi Gui, for their genuine support, guidance, patience, encouragement, and friendship. I am forever honored to have been their student. John, from the moment you interviewed me for the Ph.D. position in early 2015, I felt a great bond with you; You are the best role model that any doctoral student would love to have with your intelligence, professionalism, kindness, while always thinking outside the box; Thank you for teaching me how to turn unorganized and somewhat ambitious ideas into rigorous mathematical statements; Thank you for setting an example on how to overcome multiple challenges during a pandemic and still delivering the best class possible. Luyi, thank you for being an amazing mentor and teaching me how to approach a research project with scientific rigor, dedication and passion; I would not have been able to advance my research in the field of mechanism design without your guidance and constant support; You are also a great role model when it comes to teaching complicated mathematical subjects in the simplest and most informative way.

I thank my other committee member, Professor Shuya Yin, for her valuable suggestions and advice and evaluating my dissertation. I had the honor of being advised by Professor Yin during the first year of my doctoral program and I am grateful for her mentorship and for showing me the correct path early in my Ph.D. journey. Many thanks to my extended advancement committee, Professors Robin Keller and Vijay Vazirani for their evaluation of my dissertation proposal and directing me to the right place. Especial thanks to Robin for her tremendous support during my job search and to Vijay for sharing his knowledge and wisdom in approximation algorithms and the theoretical foundations of algorithm design.

The first chapter of this dissertation is the result of a joint work with my colleague and long–time friend, Mojtaba Hosseini, and our advisor, John Turner for which I am extremely grateful for their collaboration, dedication and for patiently hearing my stories of the basketball league and bearing with me while we were trying to develop a simple idea into a mature research project. Mojtaba and I have been together from our undergraduate program at Sharif University in Tehran. Mojtaba, thanks for being such a wonderful friend and colleague, and for always being there for me whenever I wanted fresh ears to judge my work and critique my analysis; thanks for always sharing your valuable insight and wisdom throughout our graduate studies in Turkey and in the US; I am forever grateful for having you as my dearest friend.

Especial thanks go to our wonderful ODT faculty at the Merage School of Business (Professors Carlton Scott, Robin Keller, Rick So, Shuya Yin, John Turner, Luyi Gui, and Ken Murphy) for their tremendous support and mentorship and for providing the friendliest and warmest learning environment for us Ph.D. students. I am particularly thankful for their generous advice and guidance during my job hunting days. I also thank Noel Negrete for devoting herself for the betterment of the Ph.D. program. I had the distinct honor of serving as the president of the Ph.D. students association at Merage School of Business for one academic year and I am grateful to Noel, Xuan Xie (our social chair), and Professor Terry Shevlin for their collaboration. I would like to express my most sincere gratitude towards my Masters advisor, Professor Emre Alper Yıldırım, who was the main reason I fell in love with research in the field of optimization and mathematical programming and pursued my graduate studies as Ph.D. applicant. I wish him the best of luck at University of Edinburgh. To my other professors from Sharif University of Technology in Tehran, Koç University in Istanbul, and University of California Irvine: Professors Hashem Mahlooji, Mohammad Asadi Garmaroudi, Serpil Sayin, Soheila Jorjani, Sameer Singh; thank you for all the inspiration and motivation you have given me throughout my career.

This journey would not have been possible without the incredible sacrifice my parents have shown throughout the years. Words cannot express how grateful I am for their genuine and pure love and encouragement, and the impact they had in shaping my character. My father has been my teacher in high school, but more importantly I have been and still am his student in life and he has helped me to develop my philosophy of life. My mother has endured so many hardships with an unmatched dedication to give me and my two sisters comfort, motherly love, and emotional support. I am deeply indebted to my two lovely sisters who have always cheered for my accomplishments, while I was far away from them pursuing my academic dream. I thank my father–in–law for always encouraging me in my journey as a graduate student and setting a great example by immersing himself in scientific projects even after retirement. To all my dear friends who have made an impact on my worldview and my personality in many amazing ways: Omid Hashemi, Hadi Feyzollahi, Ashkan Zakaryazad, Yahya Roshanzamir, Abbas Maazallahi; many other close friends with whom I have shared the most memorable moments in the past several years (MirAkbar S., Amir B., Iman A., Mahdi P., Ali F., Ashkan S., Saeed S., Amirhossein G., Ali E., Saman R., Mohammad T., Reza K., Alireza S., Timothy H., Trey H.); and my fellow doctoral students at the Paul Merage School of Business (Ali Hojjat, Vahid, Ali Esmaeeli, Jiaru, Yuhan, Yiwei, Mojtaba, Alex, Qi, Amy, Henry, Xianjia, Cristina, Jin Sik, Jooho, Aruhn, Luming, Rico, Sardar), I wish you all health, happiness and much success in your future endeavors.

Finally, I am eternally grateful to my soulmate and partner–in–life, Zahra Montazeri, for her unconditional love, attention, encouragement, friendship, believing in me and always pushing me to set sight on new heights regardless of the challenges ahead. Zahra, thank you for always encouraging me to try new food while I was insistently refusing; thank you for bearing with me while I was watching sports games with excitement and occasional yelling and screaming; thank you for always being a great travel companion when I made last–minute arrangements for a road trip; I have grown so much being alongside of you, and I am forever thankful for having you as my wife and my best friend; I wish you all the best in your new role in Manchester, UK.

CURRICULUM VITAE

Ali Hassanzadeh Kalshani

EDUCATION

Ph.D., Operations and Decision Technologies Sept. 2021
Paul Merage School of Business, University of California Irvine, Irvine, CA
Dissertation: "New Operations Research Models for Emerging Problems in Production, Service, and Sports"
Advisors: Dr. John Turner, Dr. Luyi Gui

M.Sc., Industrial Engineering July 2015
College of Engineering, Koç University, Istanbul, Turkey
Thesis: "Optimization-based Algorithms for the Graph Partitioning Problem"
Advisor: Dr. Emre Alper Yıldırım

B.Sc., Industrial Engineering June 2013
Department of Industrial Engineering, Sharif University of Technology, Tehran, Iran

RESEARCH INTERESTS

Applied Operations Research; Resource Sharing; Data-driven Optimization; Business Analytics; Machine Learning; Combinatorial Optimization; Graph Theory; Algorithm Design

TEACHING INTERESTS

Management Science; Operations Management; Supply Chain Management; Large-scale Optimization; Network Models and Algorithms; Predictive Analytics; Statistics; Intro to Machine Learning

RESEARCH EXPERIENCE

- **How to Conclude a Suspended Sports League?**, invited for re-submission, *Manufacturing & Service Operations Management*, (with Mojtaba Hosseini and John Turner)
- **Analysis of Pricing Mechanisms under the Resource Exchange Economy**, Working Paper (with John Turner and Luyi Gui)
- **Two Optimization-based Algorithms for the Graph Partitioning Problem**, Working Papaer (with Emre Alper Yıldırım)

- **Predicting Changes in Stock Market Indices Using Daily News Headlines**, available at SSRN (ssrn.com/abstract=3685530) (with Ahmad Razavi and Reza Asadi)

TEACHING EXPERIENCE

Recitation Leader/ Teaching Assistant:

- Management Science (undergrad & MBA core), UC Irvine – (Fall 2017-'18-'19, Winter '18-'19-'20, Spring '16-'17-'18-'19, Summer '17-'18)
- Management Science for Analytics (BANA core), UC Irvine (Winter 2021)
- Predictive Analytics (MBA & BANA elective), UC Irvine (Fall '17-'18, Spring '17-'20-'21)
- Statistics for Management (MBA core), UC Irvine (Winter 2016, Summer '16)
- Operational Excellence (MBA elective), UC Irvine (Fall '16, '20)
- Analytical Decision Making (MBA elective), UC Irvine (Winter '18-'19)
- Discrete Mathematical Structures (undergrad core), Koç University (Fall '13-'14)
- Introduction to Statistics (undergrad core), Koç University (Spring '14)
- Logistics Management (undergrad & grad elective), Koç University (Spring '15)

Teaching Workshops:

- Transition to Remote Teaching (attended 8 workshops), UC Irvine Summer 2020
- TA Professional Training and Development Program (two-day workshop), UC Irvine Sept. 2015

CONFERENCE PRESENTATIONS

Concluding a Suspended Sports League:

- INFORMS 2020 Virtual Annual Meeting (Track WC39) Nov. 2020

Pricing Mechanisms in Resource Exchange Economy:

– INFORMS 2017 Annual Meeting, Houston, TX (Track WB35) Oct. 2017

– Annual Research Festival, Paul Merage School of Business May 2018

– INFORMS 2019 Annual Meeting, Seattle, WA (Track TC55) Oct. 2019

– School of Business, Aarhus University, Aarhus, Denmark Nov. 2019

– Faculty of Engineering, Lund University, Lund, Sweden Dec. 2019

Optimization-based Algorithms for the Graph Partitioning Problem:

– Annual Research Festival, Paul Merage School of Business May 2016

– INFORMS 2017 Annual Meeting, Houston, TX (Track WC69) Oct. 2017

HONORS & AWARDS

- Division of Teaching Excellence and Innovation **(DTEI) Fellowship Award**, UC Irvine, June 2020
- **Ray Watson Doctoral Fellowship Award**, Paul Merage School of Business, UC Irvine, May 2019
- **Third Best Poster Presentation Award**, Annual Research Festival, Paul Merage School of Business, May 2018
- **Management Fellowship Award**, highest score in PhD comprehensive exam, Paul Merage School of Business, UC Irvine, Sept. 2017
- **Five-year Doctoral Fellowship**, Paul Merage School of Business, UC Irvine, Sept. 2015
- **TUBITAK Project (112M870) Scholarship**, Koç University, Istanbul, Turkey, Sept. 2013

ABSTRACT OF THE DISSERTATION

New Operations Research Models for Emerging Problems in Service Operations

By

Ali Hassanzadeh Kalshani

Doctor of Philosophy in Management

University of California, Irvine, 2021

Associate Professor Luyi Gui, Chair
Associate Professor John Turner, Chair

With the goal of exploiting theoretical and practical advancements in operations research to solve important problems in production, service, and sports, this dissertation studies two problems in particular using optimization: proposing a plan to conclude a suspended sports league in a shortened time frame, and analyzing pricing mechanisms in resource exchange economic models.

First, we study the problem of concluding a suspended sports league in a shortened time frame. Professional sports leagues may be suspended due to various reasons such as the recent COVID-19 pandemic. A critical question arises when the league decides to select a subset of the remaining games to conclude the season in a shortened time frame. Despite the rich literature on scheduling an entire season starting from a blank slate, concluding an existing season is quite different. Our approach attempts to achieve rankings similar to that which would have resulted had the season been played out in full. We propose a data-driven model which exploits predictive and prescriptive analytics to produce a schedule for the remainder of the season comprised of a subset of originally-scheduled games in anticipation of the future outcomes. This not only requires us to introduce novel rankings-based objectives, but also requires us to consider stochastic modeling approaches as well as a predictive model

for estimating the parameters in the stochastic optimization model. In comparison, all of the sports scheduling optimization models in the literature are deterministic. We study the efficacy of our approach through comprehensive computational and simulation experiments. We present simulation-based numerical experiments from previous National Basketball Association (NBA) seasons 2004–2019, and show that our models are computationally efficient and produce interpretable results. Our approach provides a data-driven decision-making framework for concluding suspended sports leagues by taking uncertainties into account. We also provide suggestions on how to conclude the 2019–20 NBA season. As an addition to this chapter, we study the problem of concluding a season after a suspension while there are no prior games played. In other words, when the hiatus happens to be at the beginning of the season, and the league starts late which makes a shortened season inevitable, a natural question is that which games or matchups should be included in the shortened season. The main challenge in this scenario is the fact that there are no prior games played in the same year and the idea of employing a predictive model does not work, unless we find a way to use past season(s) games to train a predictive model. In order to overcome this challenge, we add player–level features to the training dataset which enables us to train a predictive model using only the previous season games to predict the outcome of the new season. Once we have a predictive model, we can use a similar approach as presented at the beginning of this chapter.

Second, we study an economic system where there are multiple agents each endowed with certain amount of resources, aiming to make profit either by producing their unique product and selling it to the spot market or by trading their endowed resources to other agents. All agents use the same type of resources for production (and exchange) purposes, possibly with different usage rates. The total profit in the system depends on the allocation of resources among agents and the production quantity of all agents. This problem can be studied from two lenses: centralized and decentralized. From a central planner's perspective, it is desirable if resources are shared among all agents and they are distributed according to the production plan with the maximum profit. However, achieving this optimal distribution of resources can be challenging in practical settings since resource sharing is typically carried out in a decentralized way, i.e., each agent, independent of, or even oblivious to, other agents' decisions, determines her resource exchange quantity. In this thesis, we study how to coordinate the resource exchange decisions among decentralized agents through resource pricing approaches. Motivated by the operations of practical resource sharing alliances (e.g., capacity sharing in transportation alliances, equipment sharing in medical networks), we consider a framework where the central planner determines the prices at which the agents exchange resources, collects and fulfills agents' resource exchange requests. We assume the existence of a spot market where imbalance between resource supply and demand within the alliance can be addressed.

The objective of the agents is to maximize their own profit which is formulated as a function of the resource price, as well as production and exchange variables. Hence, the choice of resource prices influence how much resource each agent is willing to sell/buy and accordingly the overall distribution of the resources in the decentralized problem. We measure the profitability of the decentralized resource exchange system through an efficiency ratio, which is defined as the worst case ratio between the aggregate profit in the decentralized system and its centralized counterpart. An efficiency ratio of one indicates that the total profit from the centralized and decentralized problems match, and the centrally optimal distribution of resources can be attained in the decentralized system. The problem of finding the resource prices under which the efficiency ratio is maximized is called the coordination problem. Coordinating decentralized resource exchanges via pricing approaches has been studied in the literature, which largely focuses on linear pricing, i.e., a constant unit price is applied for each resource. Our work first shows that linear pricing does not guarantee an efficiency ratio of one. Nonlinear price functions studied in the literature can potentially achieve an efficiency ratio of one,, but they suffer from two drawbacks: price discrimination (i.e., different agents pay according to different pricing schemes) and the need to subsidize the exchange

transactions by the central planner. In this thesis, we study whether an efficiency ratio of one can be achieved under nonlinear pricing functions that apply the same unit price to all agents and that requires no subsidization from the central planner. We focus on a quadratic pricing function, and show that under certain conditions, (i.e., a constant term plus the product of another constant term and the exchange quantity) it achieves an efficiency ratio of one without discriminating among agents and with minimal subsidy. We finally extend our analysis to stochastic cases where the agents' revenue information is not fully–known when the resources pricing scheme is determined. We show that uncertainty undermines the effectiveness of the quadratic pricing scheme we proposed in the deterministic case in that it does not guarantee an efficiency ratio of one even under the conditions previously proposed. Nevertheless, we can numerically identify the quadratic pricing function that maximizes the efficiency ratio in the stochastic case, and we show the usefulness of our approach based on extensive numerical results.

Chapter 1

Concluding a Suspended Sports League

1.1 Introduction

The novel coronavirus named SARS-CoV-2, the causative agent of COVID-19 (Ayres 2020), emerged in December 2019, and the World Health Organization (WHO) declared the outbreak a pandemic with over 178 million individuals worldwide, and over 33 million cases in the US, infected by the disease as of June 22, 2021 (WHO 2021). The outbreak disrupted many social activities at a global scale, exposed stark problems in the healthcare systems with governments enforcing unprecedented quarantine, lockdowns, travel restrictions, and social distancing measures to reduce transmission of the virus (Dorsett 2020, Kaplan 2020). As a result of the lockdown, many entertainment industries and professional sports leagues worldwide ceased activity indefinitely. With the rapid spread of COVID-19, the National Basketball Association (NBA) was the first professional sports league in the US to suspend games, effectively pausing the season as of March 11, 2020 (NBA 2020a). The goal in this chapter is to propose a policy to conclude the suspended season.

Professional sports leagues around the world adopt various tournament formats, with round-robin and elimination as the two widely used tournament styles. In a round-robin tournament each contestant plays against all other contestants in turn. The structure of some leagues (e.g., soccer) consists only of a round-robin tournament with the highest-ranking team at the end of each season recognized as the champion. In contrast, other sports leagues (e.g., basketball, football, hockey, and baseball) have a round-robin regular season followed by an elimination tournament (postseason or playoffs) in which only a subset of teams qualify based on their regular season performance. The most popular sports associations in the US including NBA, National Football League (NFL), National Hockey League (NHL), and Major League Baseball (MLB) follow an asymmetrical round-robin structure, in that the number of games between any two teams depends on their conference and divisional affiliations, followed by playoffs. Even though our focus in this chapter is on the NBA, because of similar league structure, our results can be extended to the NFL, NHL, MLB, or any sports league with asymmetrical round-robin tournament format.

In an NBA regular season, there are a total of 1,230 games. Once the schedule is set, all teams, venues, broadcasters and the press know their schedule for the entire season, and can plan their activities accordingly. Adjustments to the schedule are required when the league is suspended for any reason, such as lockouts due to player strikes. Consequently, our methodology can also be applied to resuming a season after such a lockout. There have been four lockouts in the history of the NBA in which the league was forced to start late (e.g., December of that year) due to the expiration of the Collective Bargaining Agreement (CBA) between league owners and the National Basketball Players Association (NBPA). In two of those four lockout instances, the regular season was shortened to 50 and 66 games per team in the seasons 1998–99 and 2011–12, respectively. Apart from these four lockouts, the recent COVID-19 pandemic is the fifth league suspension in the NBA's history. Overall, suspensions are an occasional occurrence of significant consequence in major sports leagues; in Table 1.1 we tally the number of suspensions and shortened seasons for the NBA, NFL, NHL, and MLB.

League	NBA	NFL	NHL	MLB
Suspension instances	5	8	4	8
Shortened Season instances	2	2	2	3

Table 1.1: The number of disruptions in regular season games.

After the NBA's COVID-19 suspension, there has been much speculation in the media on the possible subsequent actions by the NBA, whether the 2019–20 regular season will be resumed, and how it will be concluded. The teams and players benefit financially when games are actually played, and their profit is dependent on the total number of games played. On the other hand, the league is also concerned with concluding the season fairly and in a timely manner which does not force the next season to start late or to be shortened. These objectives are mostly in agreement, but some conflicts of interest arise. The main plausible directions that the NBA can take to conclude the current season are the following:

1. The NBA cancels the regular season along with the playoffs. In this scenario, which is not likely to happen, a champion will be determined by vote.

2. The league cancels the remaining games, and the top-ranked teams as of the suspension date qualify for the playoffs. This scenario, also unlikely, would be regarded as unfair.

3. The NBA resumes the regular season with all 259 remaining games scheduled to play.

4. The NBA selects a subset of the remaining 259 games to be played (shortened season).

Options 3 and 4, which we compare extensively throughout the chapter, are illustrated in Figure 1.1. Note that the rankings of the teams (by number of wins over the played games) depends on the specific set of games that are played.

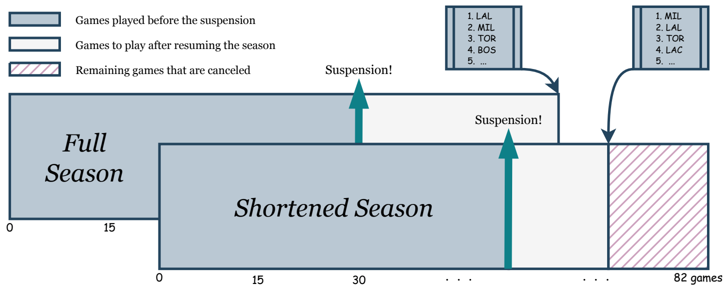


Figure 1.1: Two strategies to conclude the league: full season vs. shortened season after resuming the league

Our focus in this study is to propose a method which chooses a subset of games to conclude a shortened season that remains an asymmetrical round-robin tournament while producing end-of-season rankings that are as close as possible to the rankings that would result had the full season been played (i.e., no games cancelled). There are, of course, many considerations that come into play when constructing a sports schedule. While our approach is less detailed than constructing a full timetable which incorporates not only the set of games to play but also their sequence (and corresponding travel schedule), our model is sufficiently general to allow for any logical constraints on the subset of games chosen, which can be driven by specific practical considerations. Moreover, our focus on providing a fair final ranking is both new in the academic literature and aligns well with the league's primary concern to conclude a shortened season in an equitable yet timely fashion.

At a high level, our model selects which games to include and which ones to exclude in the remainder of the season. To make these decisions, we develop several model components. First, we use a predictive model to predict the outcomes of all games in the season that have not yet been played. Then, we use these predicted outcomes to produce a projected ranking which is our best estimate of how the season would conclude if all games were played; this is used as a target ranking. We then use a prescriptive model to select games so that, in expectation, the resulting ranking is as close to the target ranking as possible.

The organization of the rest of the chapter is as follows. We review the related works in the literature in section 1.2, followed by a description of the problem under study and background in section 1.3. In section 1.4.1 we introduce our predictive methodology and the feature selection procedures. In section 1.4.2, we lay out the prescriptive model assumptions, notation, as well as four possible choices for the objective function, each resulting in a stochastic optimization problem. In section 1.5 we introduce two approximation schemes, namely, mean value approximation and sample average approximation, as well as variable fixing techniques for solving two of our stochastic models, followed by an exact deterministic counterpart for solving the third stochastic model. Computational experiments and suggestions for concluding the 2019–20 season are presented in section 1.6, followed by discussions, limitations, and future work in section 1.7. Finally, we conclude the chapter in section 1.8.

1.2 Literature Review

A distinguishing attribute of this study is the two-phase analytics approach combining predictive and prescriptive models. In this section, we review existing literature in both directions: predicting the game outcome in sports using predictive models, and scheduling sports leagues using optimization and algorithm design techniques.

Predictive models. There is an extensive literature on predicting the outcome of games in sports. Thabtah et al. (2019) have studied three binary classifiers including Naïve Bayes, Logistic Regression, and Neural Networks using team statistics (e.g., offensive and defensive performance metrics). Igiri and Nwachukwu (2014) used the same classifiers to make predictions in soccer. Brown and Sokol (2010) combined logistic regression with a Markov chain model in predicting National Collegiate Athletic Association (NCAA) basketball games, in an attempt to capture the causal effect between games throughout the season. In a similar work, Arkes and Martinez (2011) studied the effect of momentum in the performance of different teams in the NBA. Other papers using binary classifiers to make predictions in sports leagues include (Magel and Melnykov 2014, Loeffelholz et al. 2009).

Scheduling basketball games. There is a stream of articles addressing some of the challenges in scheduling sports leagues including fairness of schedule, broadcasting, and travel distances. While some papers focus on one aspect of the scheduling problem as the main objective, others propose multi-objective models. Using optimization-based approaches in scheduling basketball games, Bean and Birge (1980) considered traveling costs and player fatigue as the main goals, Weiss (1986) studied the schedule bias between the regular season and post-season, while Westphal (2014) considered venue availability and broadcasting considerations as the main objectives. To propose a schedule for the NCAA basketball games, Nemhauser and Trick (1998) and Henz (2001) have applied integer programming and constraint programming, respectively. There are also papers that develop tailored algorithms, often based on graph theory, for scheduling basketball games including (Wright 2006, Lewis and Thompson 2011, Januario et al. 2016, Drexl and Knust 2007, Briskorn and Drexl 2009). We suggest survey papers (Rasmussen and Trick 2008, Kendall et al. 2010) for an overview of round robin scheduling studies. For a comprehensive list of articles in the broader scope of analytical methodologies applied to sports, including optimization and probabilistic modeling, see (Fry and Ohlmann 2012a,b).

Other sports. Even though the main focus in this study is the NBA, it can be extended to other sports leagues with similar asymmetric round-robin formats such as football, baseball, and hockey. In an asymmetric round-robin tournament, the number of games between two teams depends on their division affiliations, as opposed to the symmetrical tournament format where each pair of teams play exactly twice in each season. The following papers use mixed integer programming (MIP) formulations to schedule games in different leagues: Fleurent and Ferland (1993) in hockey, Kostuk and Willoughby (2012) in football, Trick et al. (2012) and Jiaqi Xu et al. (2019) in baseball. There are a stream of articles addressing the scheduling problem in symmetrical round-robin sports (e.g., soccer and volleyball), using optimization formulations, including (Durán et al. 2012, Goossens and Spieksma 2009, Cocchi et al. 2018). For examples of papers studying the same type of scheduling problem in non-sports applications see (Freeman et al. 2016, Gans et al. 2015).

Although there have been many optimization models proposed for scheduling an entire season starting from a blank slate, the problem that we consider (that of concluding an existing season) is quite different and to the best of our knowledge has not been previously studied. A significant difference is that our model attempts to achieve rankings similar to that which would have resulted had the season been played out in full; this not only requires us to introduce novel ranking-based objectives, but also requires us to consider stochastic modeling approaches as well as a predictive model for estimating the parameters in the stochastic optimization model. In comparison, all of the sports scheduling optimization models that we cite in our literature review are deterministic.

1.3 Problem

In this section, we first provide an overview of how regular season games function in the NBA and describe the main points of concern after the recent COVID-19 suspension. We then frame the conclusion of the season as a problem of selecting a subset of the remaining games, which requires us to introduce several ranking similarity metrics.

1.3.1 Background

The NBA is composed of 30 teams which are divided into two conferences of three divisions with five teams each. The list of teams, respective divisions and conferences are given in Figure 1.2. In an NBA regular season which spans approximately 180 days starting in October and finishing in April of each year, a team plays a total of 82 games, according to the following formula: four games against the other four division opponents ($4 \times 4 = 16$ games), four games against six (out-of-division) conference opponents ($4 \times 6 = 24$ games), three games against the remaining four conference teams ($3 \times 4 = 12$ games), and finally two games against teams in the opposing conference ($2 \times 15 = 30$ games). A five-year rotation determines which out-of-division conference teams are played only three times. After five seasons, each team will have played 20 games against each in-division opponent, 18 games against each out-of-division opponent, and 10 games against each team from the opposing conference.

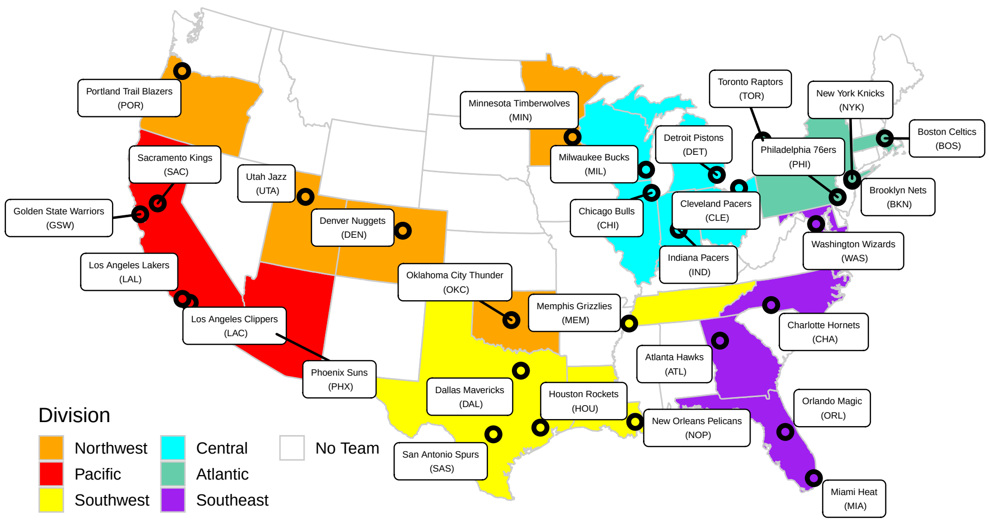


Figure 1.2: The NBA is composed of two conferences, six divisions and 30 teams. The Eastern Conference is comprised of the Central, Atlantic, and Southeast divisions, while the Western Conference consists of the Northwest, Pacific, and Southwest divisions.

At the conclusion of the regular season, the 8 top-ranked teams in each conference (16 in the league) advance to the playoffs. In each conference, the team with rank i is matched to the team with rank $8 - i$, for $i \in \{1 \dots 8\}$, and each matchup winner proceeds to the next round, with all matchups occurring within-conference until the final matchup which pits the winning team of the Eastern conference against the winning team of the Western conference. All playoff matchups are best-of-seven series, i.e., a team needs to win four out of seven games against the same opponent to win the matchup. Moreover, the highest-ranked team is given home court advantage and hosts games 1, 2, 5, and 7, while the lower-ranked team hosts games 3, 4, and 6 (with games 5–7 played only if needed). Note that due to how teams are matched, the top-ranked four teams in each conference are always given home court advantage in the playoffs.

There is a strong connection between a team's regular-season ranking and its playoff performance, which can be seen by looking at the history of 73 completed NBA seasons. More than 75% of playoff series are won by the team with home court advantage. In only five seasons did an 8^{th}-ranked team win a playoff series against a 1^{st}-ranked team. Of the 73 NBA champions, 71 were ranked among the top three teams in the league. Only three teams with winning percentage less than 0.5 have ever reached the NBA finals, and none have won the championship (NBA 2020c).

To the extent that end-of-season rankings give teams preferential treatment in the playoffs which can boost a team's chances of winning a championship, it is in the league's best interest to ensure that the ranking is fair, i.e., reflects to the greatest extent possible which teams are truly the best. Fairness can be in question when the season ends early. This is because when a specific subset of games are chosen to conclude a shortened season, it is possible for some teams to be matched with relatively easy-to-beat teams while others are matched with harder-to-beat teams, and this may result in a ranking that is quite different than one which would have resulted had the season been played in full. (We assume the full season's ranking is fair, since the league constructs the full season schedule in a balanced and equitable manner, and in general the public accepts the ranking at the end of the full season as fair).

Figure 1.3 shows the NBA ranking at the time of the 2019-20 COVID-19 suspension. There are a number of reasons to believe that the ranking resulting from playing the full season would be considerably different than the ranking in place at the time of suspension. In particular:

		Western Conference						Eastern Conference			
League	West		Wins	Losses	Win%	League	East		Wins	Losses	Win%
2	1	Los Angeles Lakers	49	14	0.778	**1**	1	Milwaukee Bucks	53	12	0.815
4	2	Los Angeles Clippers	44	20	0.688	**3**	2	Toronto Raptors	46	18	0.719
6	3	Denver Nuggets	43	22	0.662	**5**	3	Boston Celtics	43	21	0.672
7	4	Utah Jazz	41	23	0.641	**8**	4	Miami Heat	41	24	0.631
9	5	Oklahoma City Thunder	40	24	0.625	**11**	5	Indiana Pacers	39	26	0.600
10	6	Houston Rockets	40	24	0.625	**12**	6	Philadelphia 76ers	39	26	0.600
13	7	Dallas Mavericks	40	27	0.597	**15**	7	Brooklyn Nets	30	34	0.469
14	8	Memphis Grizzlies	32	33	0.492	**16**	8	Orlando Magic	30	35	0.462
17	9	Portland Trail Blazers	29	37	0.439	**22**	9	Washington Wizards	24	40	0.375
18	10	New Orleans Pelicans	28	36	0.438	**23**	10	Charlotte Hornets	23	42	0.354
19	11	Sacramento Kings	28	36	0.438	**24**	11	Chicago Bulls	22	43	0.338
20	12	San Antonio Spurs	27	36	0.429	**25**	12	New York Knicks	21	45	0.318
21	13	Phoenix Suns	26	39	0.400	**26**	13	Detroit Pistons	20	46	0.303
28	14	Minnesota Timberwolves	19	45	0.297	**27**	14	Atlanta Hawks	20	47	0.299
30	15	Golden State Warriors	15	50	0.231	**29**	15	Cleveland Cavaliers	19	46	0.292

Figure 1.3: NBA ranking at the time of suspension on March 11, 2020.

- The race for clinching the last playoff spot is very much alive in the western conference, where the teams ranked 8 through 11 have all won between 28–32 games. They are within 1–4 games of each other and any one of them could be ranked 8^{th} at the end of the season (recall the top 8 teams in each conference advance to the playoffs).

- The `Philadelphia 76ers` was considered a title contender at the beginning of the season, yet is currently ranked only 6^{th} in the eastern conference. Despite some early losses, the `76ers` are expected to bounce back and rank among the top 4 teams in the east, which would give it home court advantage in the playoffs. This is noteworthy, since the `76ers` historically perform disproportionately better at home than on the road. Their win ratio thus far in 2019-20 is 29/31 at home, while they have lost almost half of their away games.

- The `Memphis Grizzlies` (currently ranked 8^{th} in the west) have a very difficult schedule in the remainder of the season, while some teams chasing the `Grizzlies`, namely the `New Orleans Pelicans` (currently ranked 10^{th}), have a much easier remaining schedule. For this reason, ranking projections from multiple sources (ESPN 2020, Sports Illustrated 2020) claim the `Pelicans` have a higher chance of qualifying for the playoffs than the `Grizzlies`.

- The `Orlando Magic` (currently 8^{th} in the east) is expected to improve its performance in the remainder of the season, which will allow it to swap places in the ranking with the currently 7^{th}-place `Brooklyn Nets`. This will prevent the `Magic` from meeting the powerful `Milwaukee Bucks` in the first round of the playoffs (recall that the 7^{th}-ranked team is matched with the 2^{nd}-ranked team, and the 8^{th}-ranked team is matched with the 1^{st}-ranked team).

Finally, it is worth noting that the order in which teams pick rookie players in each year's NBA draft is also tied to the final ranking, with the lowest-ranked teams having a higher chance of winning the lottery and drafting the best rookie player. Therefore, the specific ranking of teams outside of the top 8 in each conference is also important. The quality of the players in the draft varies from season to season, but some first-pick rookie players have included generational talents *LeBron James*, *Magic Johnson*, and *Hakeem Olajuwon*, who have had a huge impact in leading their respective teams to win multiple championships.

1.3.2 Problem Description

At the time of the 2019-20 suspension, 971 games in the 1230-game season were played leaving 259 games remaining. Given a target number of games that each team should play in the full season, we are interested in selecting a subset of the remaining 259 games that satisfies these targets. Typically, each of the 30 teams plays 41 home and 41 away games for a total of 82 games in the season. Shortening the season involves reducing the target from 82 games/team to a lower number (e.g., 70), with half the games at home and half away. Since the results of the 259 remaining games are uncertain, both the ranking produced by playing the full 82 game/team season and the ranking produced by playing a shortened (e.g., 70 game/team) season are uncertain. Our problem is to select a subset of games that minimize the expected dissimilarity between the ranking of the full season and the ranking of the shortened season. Before we introduce our models, we introduce several metrics that may be used for measuring the similarity of rankings.

1.3.3 Measures of Similarity/Dissimilarity between Rankings

We represent a ranking of n teams as a vector, with the components $1..n$ permuted in order of highest-to-lowest percentage of games won during the regular season. Throughout the chapter, we follow the convention that $\hat{r}$ represents a ranking resulting from playing all games in the full season, while r represents a ranking resulting from playing a specific subset of the remaining games (i.e., the shortened season case). Furthermore, when we wish to notationally distinguish between multiple rankings in the shortened season case, we use a superscript. For example, $r^{(1)}$ and $r^{(2)}$ represent two distinct rankings resulting from concluding a shortened season with two different sets of games.

Three widely used measures of similarity/dissimilarity between rankings are *Kendall's coefficient* (τ), *Spearman's coefficient* (ρ), and *Manhattan distance*. We will now define these metrics, using the following small 4-team example to illustrate.

Example 1: Assume there are only four teams in the league: `LAL`, `MIL`, `LAC`, and `BOS`. Table 1.2 contains the full-season ranking $\hat{r}$, and two alternative rankings $r^{(1)}$ and $r^{(2)}$ (which is simply the reverse of $r^{(1)}$).

Teams	Ranking ($\hat{r}$)	Ranking ($r^{(1)}$)	Ranking ($r^{(2)}$)
LAL	1	1	4
BOS	2	4	1
MIL	3	2	3
LAC	4	3	2

Table 1.2: Example 1: Three different rankings.

1.3.3.1 Kendall's τ.

Kendall's rank correlation coefficient, or simply Kendall's τ, introduced in Kendall (1938), is a metric used to measure the ordinal association between two measured quantities. Intuitively, Kendall's τ is high (maximum +1) when observations in two variables have similar ranks, and it is low (minimum -1) when observations have dissimilar (opposite) ranks. Let us consider Kendall's τ in the context of team rankings in the NBA. For a given pair of rankings $(r, \hat{r})$, we call a pair of teams (i, j) *concordant* if the preference between these two teams agrees in both rankings; that is, if both $\hat{r}_i > \hat{r}_j$ and $r_i > r_j$, or if both $\hat{r}_i < \hat{r}_j$ and $r_i < r_j$. The pair of teams (i, j) is said to be *discordant* if $\hat{r}_i > \hat{r}_j$ and $r_i < r_j$, or if $\hat{r}_i < \hat{r}_j$ and $r_i > r_j$. If $\hat{r}_i = \hat{r}_j$ or $r_i = r_j$, the pair of teams is neither concordant nor discordant. For instance in Table 1.2, when we compare rankings $\hat{r}$ and $r^{(1)}$, the pair (`LAL`, `BOS`) is concordant, since in both rankings `LAL` stands higher than `BOS`. The pair (`BOS`, `LAC`) however is discordant, since `BOS` has the higher rank in $\hat{r}$, while `LAC` gets the higher spot in $r^{(1)}$.

Using the number of concordant and discordant pairs and the total number of teams in the ranking, Kendall's τ is defined as

$$\tau_{CD} = \frac{(\text{number of concordant pairs}) - (\text{number of discordant pairs})}{\binom{n}{2}} \tag{1.1}$$

For the rankings presented in Example 1, $\tau_{CD}(r^{(1)}, \hat{r}) = \frac{4-2}{6} = 0.33$, because there are 6 pairs of teams in total, two of which are discordant (i.e., `MIL`–`BOS`, `LAC`–`BOS`), and the other four pairs are concordant. Similarly, $\tau_{CD}(r^{(2)}, \hat{r}) = -0.33$ since the ranking $r^{(2)}$ is the reverse of $r^{(1)}$. Because $r^{(1)}$ has a higher Kendall's τ than $r^{(2)}$, it more closely resembles the full-season ranking $\hat{r}$.

Since our goal is to propose a ranking that best resembles the end-of-season ranking, $\hat{r}$, we count the pairs of teams with identical positions in $\hat{r}$ and r as concordant pairs. Consequently, provided that the number of discordant and concordant pairs add up to the total number of pairs of teams in the league (i.e., $\frac{1}{2}n(n-1)$), we redefine Kendall's τ so that it depends only on the number of concordances. Moreover, for the sake of interpretability, we also divide the number of concordant pairs (the maximum is $\frac{n}{2}(n-1)$) by $\frac{n}{2}$ to obtain the *concordance per team*, which results in a number between 0 and $n-1$. Concordance per team represents the number of opponents with the same position in two alternative rankings relative to an arbitrarily chosen team, and is defined as:

$$\tau_C = \frac{\text{number of concordant pairs}}{n/2}. \tag{1.2}$$

Continuing our example, the concordance per team is $\tau_C(r^{(1)}, \hat{r}) = \frac{4}{2} = 2$ between $(r^{(1}, \hat{r})$ and it is $\tau_C(r^{(2)}, \hat{r}) = \frac{2}{2} = 1$ between $(r^{(2}, \hat{r})$. Also note that, in general, $\tau_{CD} = \frac{2}{n-1}\tau_C - 1$.

1.3.3.2 Euclidean distance (Spearman's ρ).

Spearman's rank correlation coefficient, or simply Spearman's ρ, introduced in Spearman (1904), is another measure of rank correlation, which is high (maximum +1) when observations in two variables have similar ranks, and it is low (minimum -1) when observations have opposite ranks. Spearman's ρ takes the Euclidean distance between two rankings and transforms it into a value between -1 and 1 using the following equation

$$\rho(r, \hat{r}) = 1 - \frac{6}{n^3 - n} \sum_{i=1}^{n} (r_i - \hat{r}_i)^2. \tag{1.3}$$

It can be easily verified that $\rho(\hat{r}, r^{(1)}) = 0.4$ and $\rho(\hat{r}, r^{(2)}) = -0.4$, which is consistent with our conclusion based on Kendall's τ that $r^{(1)}$ is the most similar to $\hat{r}$.

1.3.3.3 Manhattan distance.

Similar to Spearman's ρ, the Manhattan distance (or ℓ_1-distance) between two rankings r and $\hat{r}$ measures the absolute value of the differences in these rankings. As in (1.2), we divide the Manhattan distance by n, which gives us the average number of places an arbitrarily-chosen team switches between the two rankings.

$$\text{Manhattan Distance}(r, \hat{r}) = \frac{1}{n} \sum_{i=1}^{n} |r_i - \hat{r}_i| \tag{1.4}$$

The Manhattan distances between $(r^{(1)}, \hat{r})$ and $(r^{(2)}, \hat{r})$ in Example 1 are 4 and 5, respectively.

1.3.3.4 Practical implications.

We now present an example to illustrate the magnitudes of our similarity/dissimilarity measures that we observe when using real data. Our example, shown in Table 1.3, is based on several anticipated changes to the ranking that pundits believe would occur between the suspension date of the 2019–20 NBA season and the end of the season assuming it is played in full. We have omitted Spearman's ρ from the table, as we have found it a difficult metric to optimize directly and therefore we have based our numerical results on concordance (as measured by the concordance-per-team metric, (1.2)) and Manhattan distance (1.4). Nevertheless, a model we introduce later in the chapter refers to Spearman's ρ in its derivation, and it is worth pointing out that the Manhattan distance metric is also similar in spirit to Spearman's ρ. Finally, given there are $n = 30$ teams, the maximum concordance is $n - 1 = 29$.

Anticipated changes in the ranking	Concordance	Manhattan
The Pelicans, with improved roster after *Zion Williamson* returns from his injury, replace the 8^{th}-ranked Grizzlies to clinch the last playoff spot in the west.	**28.73**	**0.26**
The 76ers move up to replace the 4^{th}-ranked Heat in the east, recovering after a bad performance for two months due to multiple injuries.	**28.73**	**0.26**
The Hawks outperform the Knicks and the Pistons in the remainder of the season, and move up to 12^{th} spot in the east, after improving their roster and acquiring *Clint Capela* and *Nene Hilario*.	**28.86**	**0.13**
The Magic steal the 7^{th}-ranked spot in the east from the Nets, who lost their two key players, *Kevin Durant* and *Kyrie Irving*, due to injury before the suspension and they will miss the remainder of the season.	**28.93**	**0.06**
All of the above changes combined	**28.13**	**0.73**

Table 1.3: Anticipated changes in the 2019–20 ranking with corresponding similarity/dissimilarity values considering rankings before and after each change.

1.4 Models

As schematically depicted in Figure 1.4, our modeling approach consists of two phases. To determine the best subset of games to include in the shortened season, we need to have estimates of the outcomes of each of the remaining games. Hence, in the first phase, we develop predictive models to predict which teams win the remaining games. Using historical data from all games in the regular season that were played before the suspension, we train a binary classification model to predict the outcome of each remaining game. Instead of including team labels while training the predictive model, we focus on game-related features (e.g., win percentage, point differential, and home-away indicator, among others). This way, we take into account the toughness of the schedule for different teams throughout the regular season.

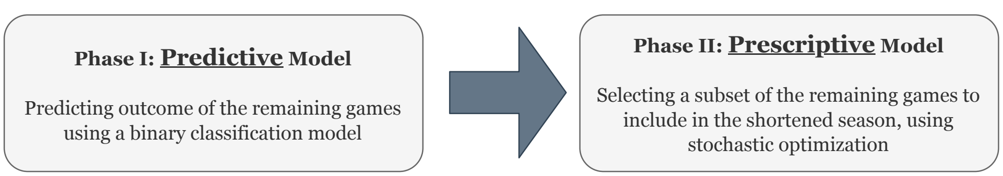


Figure 1.4: Two main phases of our methodology

Our goal in the second phase (i.e., prescriptive model) is to determine which games to include and which games to cancel in the shortened season. We minimize the expected dissimilarity between the shortened season's ranking and the full season's ranking, where this expectation is taken over multiple possible scenarios which reflect the random chance each team has for winning each game. Specifically, we treat the outcomes of games as Bernoulli random variables whose parameters are estimated in the first (predictive modeling) phase, and formulate our prescriptive models as stochastic optimization problems.

1.4.1 Predictive Model

In this section, we present a model for the probabilistic outcome of the games which are postponed due to the suspension. Since the response variable (i.e., whether the home team wins or loses) is binary, the prediction problem is a *binary classification* task. In a binary classification task, one is interested in separating a set of data points G with class labels 0 or 1. Each training data point $g \in G$ is represented by $(\vec{x}_g, y_g)$ with $\vec{x}_g = (x_{g,1}, \ldots, x_{g,D})$ and $y_g \in \{0, 1\}$ denoting the features and the class label of the data point, respectively, where D is the dimension of the feature space. The goal of the classification problem is to learn a discrimination rule $p : \mathbb{R}^D \to [0, 1]$, which represents the probability that the data point belongs to class 1 ("home team wins"). The class label of observation $\vec{x}$ is then determined by comparing $p(\vec{x})$ with a predefined threshold (e.g., 0.5). A large number of explanatory variables are available in the dataset among which we choose the most significant features for the predictive model, listed in Table 1.4.

Features	**Definition**
x_1 (resp. x_5)	home (resp. guest) team win percentage
x_2 (resp. x_6)	home (resp. guest) team average point differential
x_3 (resp. x_7)	home (resp. guest) team win percentage in the last 8 games
x_4 (resp. x_8)	home (resp. guest) team win percentage at home (resp. as guest)

Table 1.4: Features used in our binary classification method.

As illustrated in Figure 1.5, the suspension day splits the entire set of games in an NBA regular season into training and test datasets.

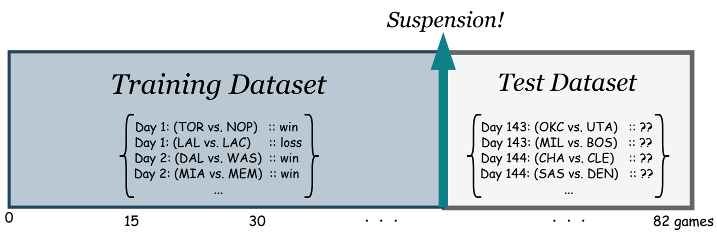


Figure 1.5: Training/test datasets in an NBA regular season

We use two metrics to evaluate the predictive performance of a classifier: *accuracy* and *predictive power*. Let G be the set of remaining games, y_g the true outcome of game g, p_g the predicted probability that the home team wins game g, and $\hat{y}_g = \mathbb{I}(p_g \geq 0.5)$ the predicted class label (i.e. most likely outcome) for game g, where $\mathbb{I}(\cdot)$ is the indicator function. Accuracy, defined below in (1.5), is a commonly-used metric for assessing the performance of a predictive model (Tharwat 2018); in our context, it measures the fraction of games that are correctly predicted according to the most likely outcome $\hat{y}_g$.

$$\text{Accuracy} = \frac{1}{|G|} \sum_{g \in G} \Big(y_g \times \hat{y}_g + (1 - y_g) \times (1 - \hat{y}_g) \Big) \tag{1.5}$$

Since we are less interested in making a single best prediction, and are more interested in solving a prescriptive model based on our predictions, we introduce the *predictive power metric* which is a modification of the commonly-used accuracy metric and we expect has independent value apart from the application in this study:

$$\text{Predictive Power} = \frac{1}{|G|} \sum_{g \in G} \Big(y_g \times p_g + (1 - y_g) \times (1 - p_g) \Big). \tag{1.6}$$

The superiority of the predictive power metric in selecting a predictive model becomes clear when the prescriptive model that uses its predictions is a stochastic optimization model. In this case, we may generate several scenarios, and consequently each game may have several predictions (one for each scenario). What becomes relevant in this context is the *average* predictive accuracy over generated scenarios (each gets treated like an independent prediction), rather than the predictive accuracy of a single "best bet". Formally, we model the outcome of each game g as a Bernoulli random event with probability p_g determined by our predictive model. As presented in Proposition 1.1 below, predictive power measures the *expected prediction accuracy* of a classifier, when these Bernoulli random processes are replicated infinitely many times.

Proposition 1.1. *Predictive power as defined in* (1.6) *measures the expected accuracy provided that the outcome of game $g \in G$, denoted W_g, follows a Bernoulli distribution with probability p_g.*

Proof. Proof. See Appendix A.1 for the proof. □

We examine five different binary classifiers including Logistic Regression (Logit), Gaussian Naïve Bayes (NB), Support Vector Machine with a linear kernel (SVM), Random Forest (RF), and Multilayer Perceptron (MLP). We provide results in section 1.6.2.

1.4.2 Prescriptive Models

In a league with n teams, let T denote the set of teams in the league. Assume that at the time of suspension, a set G of regular-season games remain to be played, and each team $i \in T$ has won a total of y_i^0 games before the suspension. We represent each game $g \in G$ with a tuple $g = (i, j, k)$, where $i(g) \in T$ and $j(g) \in T$ denote the host and guest teams, respectively, and $k(g)$ the k^{th} match between these two teams (recall the same pair of teams may play each other more than once in the season). We also define $G_i^h \subset G$ and $G_i^a \subset G$ as the set of remaining home and away games for team i, respectively.

We model the outcome of game g using the Bernoulli random variable W_g, which is one if the host team $i(g)$ wins, and zero if the guest team $j(g)$ wins. For each game g, we estimate the parameter $p_g = P(W_g = 1)$ using historical data as discussed in section 1.4.1. Formally, we also denote the set of all possible outcomes of all games in the remainder of the full season by Ξ, and use $\xi \in \Xi$ to index a specific realization of all games' outcomes. When appropriate, we will explicitly write $W_g(\xi)$ to indicate W_g's dependence on ξ. For a given outcome $\xi \in \Xi$, the total number of wins for team i after playing all remaining games (i.e., at the end of the full regular season) is $\hat{y}_i(\xi)$, where

$$\hat{y}_i(\xi) = y_i^0 + \sum_{g \in G_i^h} W_g(\xi) + \sum_{g \in G_i^a} (1 - W_g(\xi)). \tag{1.7}$$

We continue to use the caret ($\hat{}$) to denote quantities which correspond to the full regular season.

For each game $g \in G$, we define a binary decision variable x_g which takes the value of one if we choose to include this game in the shortened season, and zero otherwise.

Note that these x-variables must be made prior to us knowing the realization of ξ. We define X as the set of feasible solutions, i.e., restrictions placed on the x-variables expressed by tactical considerations such as having an equal total number of home/away games for each team as well as the integrality requirements on the x-variables, i.e.,

$$X = \left\{ \begin{array}{ll} \sum_{g \in G_i^h} x_g = m_i^h, & \forall i \in T \\ \sum_{g \in G_i^a} x_g = m_i^a, & \forall i \in T \\ x_g \in \{0, 1\}, & \forall g \in G \end{array} \right\}, \tag{1.8}$$

where m_i^h and m_i^a denote the targeted number of home and away games for team i, respectively. For instance, if team i has played 33 home and 31 away games so far before the suspension, and we decide to conclude the season with a total of 72 games for each team, then this team must play an additional $m_i^h = \frac{72}{2} - 33 = 3$ home and $m_i^a = \frac{72}{2} - 31 = 5$ away games. Another alternative would be to combine the constraints on the number of home/away games for each team into a single constraint of the form $\sum_{g \in G_i^h \cup G_i^a} x_g = m_i^h + m_i^a$ that sets a target for the total number of games to play without specific home/away sub-targets.

For a given shortened season $x \in X$ and realization $\xi \in \Xi$, we denote the total number of wins for team i at the end of the shortened regular season as $y_i(x, \xi)$, where

$$y_i(x, \xi) = y_i^0 + \sum_{g \in G_i^h} W_g(\xi) x_g + \sum_{g \in G_i^a} (1 - W_g(\xi)) x_g. \tag{1.9}$$

Let $d(y(x, \xi), \hat{y}(\xi))$ be a measure of dissimilarity between the vectors $y(x, \xi)$ and $\hat{y}(\xi)$, i.e., a measure that compares the wins accumulated by each team over a full season with the wins accumulated by each team over a shortened season, for a specific shortened season x and outcome ξ. Note that there is a one-to-one correspondence between $\hat{y}(\xi)$ and the team rankings at the end of the full season, and between $y(x, \xi)$ and team rankings at the end of the shortened season. Therefore, $d(y(x, \xi), \hat{y}(\xi))$ can also be viewed as a dissimilarity measure between these rankings, and our goal is to find a shortened season x that minimizes the expected value of this dissimilarity. That is, we are interested in solving stochastic optimization problems of the general form

$$\min_{x \in X} \quad \mathbb{E}_\xi[d(y(x, \xi), \hat{y}(\xi))],$$

for different choices of the dissimilarity measure d. We now introduce several such formulations.

1.4.2.1 Maximizing concordance per team.

For a given outcome ξ, let $\hat{r}(\xi)$ denote the ranking vector we get when the full season is played, and $r(x,\xi)$ denote the ranking vector we get when the shortened season x is played. To maximize the expected similarity between $r(x,\xi)$ and $\hat{r}(\xi)$ according to the average concordance per team metric as defined in (1.2), we solve the following stochastic optimization problem:

$$\max_{x \in X} \quad \mathbb{E}_\xi \left[\tau_C \left(r(x,\xi), \hat{r}(\xi) \right) \right]. \tag{1.10}$$

While this formulation is compact, its objective function is highly nonlinear; consequently, we linearize it as follows. First, we define a parameter $\hat{z}_{ij}(\xi)$ which takes value one if team i is above team j in the full-season ranking $\hat{r}(\xi)$, and zero otherwise. Similarly, we introduce a binary variable $z_{ij}(x,\xi)$ which takes value one if team i is above team j in the shortened-season ranking $r(x,\xi)$, and zero otherwise. Since $z_{ij}(x,\xi) + z_{ji}(x,\xi) = 1$, we introduce only the z_{ij}-variables where $i < j$ and use $1 - z_{ij}(x,\xi)$ in place of $z_{ji}(x,\xi)$ whenever it is needed. As well, we introduce continuous variables $y_i(x,\xi)$, $i \in T$, to keep track of the number of wins team i makes in the shortened season x under realization ξ. Finally, since it is clear that the solution to this optimization problem encodes a single x-vector, we henceforth suppress the x-argument for the y- and z-variables. Using these parameters and variables, we restate problem (1.10) as the following stochastic MIP:

$$[\text{PC}] \;\; \max \quad \frac{2}{n}\, \mathbb{E}_\xi \left[\sum_{i \in T} \sum_{j \in T : j > i} \left(z_{ij}(\xi)\hat{z}_{ij}(\xi) + (1 - z_{ij}(\xi))(1 - \hat{z}_{ij}(\xi)) \right) \right] \tag{1.11}$$

$$\text{s.t.} \quad y_i(\xi) = y_i^0 + \sum_{g \in G_i^h} W_g(\xi) x_g + \sum_{g \in G_i^a} (1 - W_g(\xi)) x_g \qquad \forall i \in T, \forall \xi \in \Xi \tag{1.12}$$

$$-\underline{M}_{ij}(\xi)(1 - z_{ij}(\xi)) \leq y_i(\xi) - y_j(\xi) \leq \overline{M}_{ij}(\xi) z_{ij}(\xi) \qquad \forall i, j \in T : i < j, \forall \xi \in \Xi \tag{1.13}$$

$$z_{ij}(\xi) \in \{0, 1\} \qquad \forall i, j \in T : i < j, \forall \xi \in \Xi \tag{1.14}$$

$$x \in X. \tag{1.15}$$

The objective function (1.11) counts the expected number of concordant pairs per team. Constraint (1.12) counts the number of wins for each team under each realization as defined by equation (1.9), and constraint (1.13) establishes the relationship between the number of wins and relative positions of teams, in which $\underline{M}_{ij}(\xi)$ and $\overline{M}_{ij}(\xi)$ are parameters with sufficiently large values (we show how to compute approproate values in section 1.5.1.3).

We remark that this formulation may be viewed as a stochastic program with recourse, where the x-variables are first-stage variables (for which there is only one choice to be made) and the y- and z-variables are second-stage "recourse" variables (for which there is one such variable for each possible outcome ξ). Note, however, that in our application there is no true recourse. Rather, y and z are auxiliary variables whose purpose is to linearize the objective function.

1.4.2.2 Minimizing Euclidean distance between rankings.

As introduced in section 1.3.3.2, Euclidean distance between rankings is a metric for measuring dissimilarity of two rankings. To find a ranking $r(x, \xi)$ that best resembles $\hat{r}(\xi)$ according to this measure, one needs to select a solution $x \in X$ that minimizes $\sum_i (r_i(x, \xi) - \hat{r}_i(\xi))^2$. Thus, we select the subset of games to be played after the suspension by solving the following stochastic optimization problem:

$$\min_{x \in X} \quad \mathbb{E}_\xi \left[\sum_i (r_i(x,\xi) - \hat{r}_i(\xi))^2 \right]. \tag{1.16}$$

Measuring the dissimilarity between two rankings using their Euclidean distance poses two challenges. First, it is difficult to analytically compute the expected value in (1.16) based on the ranking differences. Second, computing $r_i(x,\xi)$ for a given shortened season x and realization ξ requires a large number of binary variables $z_{ij}(\xi)$, the linking constraints (1.13), and constraints of the form

$$r_i(x,\xi) = n - \sum_{j>i} z_{ij}(x,\xi) - \sum_{j<i} (1 - z_{ji}(x,\xi)) \qquad \forall i \in T, \tag{1.17}$$

which conjointly with the quadratic objective function in (1.16), make solving even moderately-sized instances of the problem challenging. To circumvent these difficulties, we propose two alternative representations for problem (1.16) in the following two subsections.

1.4.2.3 Minimizing Manhattan distance between rankings.

In the first alternative measure of dissimilarity, we replace the Euclidean distance between rankings with *Manhattan distance* to cast the problem as a simpler stochastic MIP of the following form:

$$\text{[PM]} \ \min \quad \frac{1}{n} \mathbb{E}_\xi \left[\sum_{i \in T} d_i(\xi) \right] \tag{1.18}$$

$$\text{s.t.} \quad y_i(\xi) = y_i^0 + \sum_{g \in G_i^h} W_g(\xi) x_g + \sum_{g \in G_i^a} (1 - W_g(\xi)) x_g \qquad \forall i \in T, \forall \xi \in \Xi \tag{1.19}$$

$$- \underline{M}_{ij}(\xi)(1 - z_{ij}(\xi)) \leq y_i(\xi) - y_j(\xi) \leq \overline{M}_{ij}(\xi) z_{ij}(\xi) \quad \forall i, j \in T : i < j, \forall \xi \in \Xi \tag{1.20}$$

$$d_i(\xi) \geq n - \sum_{j>i} z_{ij}(\xi) - \sum_{j<i}(1 - z_{ji}(\xi)) - \hat{r}_i(\xi) \quad \forall i \in T, \forall \xi \in \Xi \tag{1.21}$$

$$d_i(\xi) \geq \hat{r}_i(\xi) - (n - \sum_{j>i} z_{ij}(\xi) - \sum_{j<i}(1 - z_{ji}(\xi))) \quad \forall i \in T, \forall \xi \in \Xi \tag{1.22}$$

$$z_{ij}(\xi) \in \{0, 1\} \quad \forall i, j \in T : i < j, \forall \xi \in \Xi \tag{1.23}$$

$$x \in X. \tag{1.24}$$

The objective function (1.18) minimizes the expected Manhattan distance between the full-season and shortened-season ranking, i.e., constraints (1.21) and (1.22) imply $d_i = |r_i(x, \xi) - \hat{r}_i(\xi)|$. All other constraints and parameters are the same as in PC, with the sole addition of the $\hat{r}_i(\xi)$ parameters.

Finally, we also note the following connections between PC and PM. First, the Manhattan distance between two rankings is zero if and only if the rankings are fully concordant. Second, the following proposition establishes another interesting connection between these two objectives:

Proposition 1.2. *For any shortened season $x \in X$ and realization $\xi \in \Xi$, let $\varphi_C(x, \xi)$ and $\varphi_M(x, \xi)$ be the objective values of PC and PM, respectively. The following relationship holds:*

$$\varphi_M(x, \xi) \leq (n - 1) - \varphi_C(x, \xi). \tag{1.25}$$

Proof. Proof. See Appendix A.1. □

As a result of Proposition 1.2, $(n-1)-\mathbb{E}_\xi[\varphi_C(x,\xi)]$ is an upper bound on $\mathbb{E}_\xi[\varphi_M(x,\xi)]$, which means that maximizing $\mathbb{E}_\xi[\varphi_C(x,\xi)]$ effectively minimizes $\mathbb{E}_\xi[\varphi_M(x,\xi)]$. But the opposite does not necessarily hold.

1.4.2.4 Minimizing win percentage distance.

A second alternative objective which is similar in spirit to minimizing the expected Euclidean distance between rankings keeps the quadratic nature of the objective but replaces the ranking vectors with the win count vectors. That is, we consider the Euclidean distance between the number of wins at the end of the shortened season $y(x,\xi)$ and the number of wins at the end of the full season $\hat{y}(\xi)$. Since the number of games played is different in the shortened and full seasons, we define the dissimilarity between $y(x,\xi)$ and $\hat{y}(\xi)$ in terms of win percentage as

$$\sum_{i\in T}\left(\frac{y_i(x,\xi)}{m}-\frac{\hat{y}_i(\xi)}{\hat{m}}\right)^2,$$

where m is the target total number of games for each team in the shortened season (e.g., 70), and $\hat{m}$ is the number of games played by each team in the full season (e.g., 82). To minimize this dissimilarity measure, we solve the following stochastic mixed integer quadratic program:

$$\begin{aligned}
\text{[PW]}\ \min\quad & \mathbb{E}_\xi\left[\sum_{i\in T}\left(\frac{y_i(\xi)}{m}-\frac{\hat{y}_i(\xi)}{\hat{m}}\right)^2\right] && (1.26)\\
\text{s.t.}\quad & y_i(\xi)=y_i^0+\sum_{g\in G_i^h}W_g(\xi)x_g+\sum_{g\in G_i^a}(1-W_g(\xi))x_g \qquad \forall i\in T,\forall\xi\in\Xi && (1.27)\\
& x\in X. && (1.28)
\end{aligned}$$

Note that this formulation does not require the binary variables $z_{ij}(\xi)$ and the associated linking constraints, making it a lighter formulation than PM. Moreover, as we shall show in section 1.5.2, unlike for PM, we may derive a closed-form expression for the expected value in the objective function (1.26), which results in a much simpler deterministic equivalent problem, despite the objective being quadratic rather than linear. On the other hand, as possible downside of this formulation, it does not model rankings directly, which are more closely tied to league outcomes than win percentages.

1.5 Solution Methodology

The stochastic optimization problems introduced in section 1.4.2 contain $2^{|G|}$ realizations of ξ, each with their own set of second-stage decision variables and constraints. As the full stochastic optimization problems are too large to solve directly, we introduce two methods which approximately solve PC and PM, as well as an exact deterministic counterpart for PW.

1.5.1 Solution Methods for the Stochastic Models PC & PM

For PC and PM, we approximately solve the stochastic optimization problems using (i) mean value approximation, and (ii) sample average approximation. Moreover, we introduce a variable fixing technique which accelerates our solution methods.

1.5.1.1 Mean value approximation.

Replacing all random parameters in a stochastic optimization problem by their expected values yields a deterministic problem known as the Mean Value Problem (MVP). In our case, we may produce MVP's for PC and PM by replacing the random variables W_g which represent the outcome of each game g with their means $p_g = \mathbb{E}[W_g]$. The y and z variables are then interpreted as expected values over all outcomes $\xi \in \Xi$, given the shortened season x. The MVP corresponding to PC is:

$$
\begin{aligned}
\text{[PC-MVP]} \ \max \quad & \frac{2}{n} \sum_{i \in T} \sum_{j \in T: j > i} (z_{ij} \hat{z}_{ij} + (1 - z_{ij})(1 - \hat{z}_{ij})) && && (1.29) \\
\text{s.t.} \quad & y_i = y_i^0 + \sum_{g \in G_i^h} p_g x_g + \sum_{g \in G_i^a} (1 - p_g) x_g && \forall i \in T && (1.30) \\
& -\underline{M}_{ij}(1 - z_{ij}) \leq y_i - y_j \leq \overline{M}_{ij} z_{ij} && \forall i, j \in T : i < j && (1.31) \\
& z_{ij} \in \{0, 1\} && \forall i, j \in T : i < j && (1.32) \\
& x \in X. && && (1.33)
\end{aligned}
$$

The MVP counterpart of the stochastic optimization problem PM, which we denote PM-MVP, can be obtained similarly.

1.5.1.2 Sample average approximation.

Sample Average Approximation (SAA) is a Monte Carlo simulation-based technique for approximating stochastic optimization problems (Kleywegt et al. 2002), which has been a prominent technique widely used in various applications (Begen et al. 2012, Gans et al. 2015, Freeman et al. 2016, Lu et al. 2018). Let $\mathcal{S} = \{\xi^{(1)}, \xi^{(2)}, \ldots, \xi^{(|\mathcal{S}|)}\}$ be an independently and identically distributed random sample of ξ. SAA reduces the size of the problem by approximating the expected value in the objective function with the sample average function. In the following, we use the superscript s to reference the second-stage variables and random parameters under scenario $s \in \mathcal{S}$. For instance, under scenario s, $W_g^{(s)}$ refers to outcome of game g, $\hat{y}_i^{(s)}$ refers to the total number of wins for team i at the end of the full season, and $y_i^{(s)}$ refers to the decision variable counting the total number of wins for team i at the end of the shortened season. We construct the SAA counterpart of the stochastic program PC by replacing the full set of outcomes Ξ with the sample set $\mathcal{S}$. The SAA counterpart of PM, denoted PM-SAA, can be formulated similarly.

$$\text{[PC-SAA]} \ \max \quad \frac{1}{|\mathcal{S}|}\sum_{s\in\mathcal{S}}\frac{2}{n}\sum_{i\in T}\sum_{j\in T:j>i}\left(z_{ij}^{(s)}\hat{z}_{ij}^{(s)} + (1-z_{ij}^{(s)})(1-\hat{z}_{ij}^{(s)})\right) \tag{1.34}$$

$$\text{s.t.} \quad y_i^{(s)} = y_i^0 + \sum_{g\in G_i^h} W_g^{(s)}x_g + \sum_{g\in G_i^a}(1-W_g^{(s)})x_g \qquad \forall i\in T, \forall s\in\mathcal{S} \tag{1.35}$$

$$-\underline{M}_{ij}^{(s)}(1-z_{ij}^{(s)}) \le y_i^{(s)} - y_j^{(s)} \le \overline{M}_{ij}^{(s)} z_{ij}^{(s)} \qquad \forall i,j\in T: i<j, \forall s\in\mathcal{S} \tag{1.36}$$

$$z_{ij}^{(s)} \in \{0,1\} \qquad \forall i,j\in T: i<j, \forall s\in\mathcal{S} \tag{1.37}$$

$$x\in X. \tag{1.38}$$

As the sample size increases, the optimal solution and the optimal value of the SAA problems converge to their 'true' stochastic counterparts with probability one (Kleywegt et al. 2002). It should also be noted that, since the outcome of the games are assumed to be independent Bernoulli random variables, in the SAA procedure one is effectively sampling from the individual Bernoulli distributions. Hence, although the total number of possible realizations is $2^{|G|}$, a significantly smaller sample size $|\mathcal{S}|$ should suffice for estimating the sample proportion with high confidence.

1.5.1.3 Variable fixing and preprocessing.

We may improve the computational efficiency of both the SAA and MVP counterparts of PC and PM by fixing certain variables at their optimal values and eliminating redundant constraints, as described by the following proposition.

Proposition 1.3. *Let $\tilde{\xi}$ be an arbitrary realization or expected value of ξ. For each team i,* *sort G_i^h and G_i^a in the non-decreasing order of $W(\tilde{\xi})$. Let U_i^h and L_i^h be the summation of $W_g(\tilde{\xi})$ values corresponding to the first and last m_i^h games in G_i^h, respectively. Similarly, let U_i^a and L_i^a be the summation of $W_g(\tilde{\xi})$ values corresponding to the first and last m_i^a games in G_i^a, respectively. Define $y_i^U = y_i^0 + U_i^h + (m_i^a - L_i^a)$ and $y_i^L = y_i^0 + L_i^h + (m_i^a - U_i^a)$ to be the optimistic and pessimistic number of wins for team i under $\tilde{\xi}$, respectively. For each pair of teams (i, j), define $M_{ij}^{max} = y_i^U - y_j^L$ and $M_{ij}^{min} = y_i^L - y_j^U$. Then:*

(i) If $M_{ij}^{max} < 0$, then $z_{ij}(\tilde{\xi}) = 0$, and the corresponding linking constraints are redundant.

(ii) If $M_{ij}^{min} > 0$, then $z_{ij}(\tilde{\xi}) = 1$, and the corresponding linking constraints are redundant.

(iii) Otherwise, $\underline{M}_{ij}(\tilde{\xi}) = -M_{ij}^{min}$ and $\overline{M}_{ij}(\tilde{\xi}) = M_{ij}^{max}$ serve as the big-M values in the linking constraints.

Proof. The statements follow from the definition of optimistic and pessimistic number of wins (i.e., y_i^U and y_i^L) for each team and definition of the z-variables. □

1.5.2 Exact Solution Method for the Stochastic Model PW

We now show how the stochastic problem PW from §1.4.2.4 can be solved using an equivalent deterministic problem. We will use the notation $\mathbb{V}$ to refer to the variance of a random variable.

Theorem 1.1. *The stochastic model PW can be solved using the following equivalent deterministic linearly-constrained quadratic mixed-integer optimization problem:*

$$\text{[PW-DQIP]} \ \min \quad \sum_{i \in T} \left(\frac{1}{m^2}(v_i + \mu_i^2) + \frac{1}{\hat{m}^2}(\hat{v}_i + \hat{\mu}_i^2) - \frac{2}{m\hat{m}}(v_i + \mu_i \hat{\mu}_i) \right) \tag{1.39}$$

$$s.t. \quad \mu_i = y_i^0 + \sum_{g \in G_i^h} p_g x_g + \sum_{g \in G_i^a} (1 - p_g) x_g \tag{1.40}$$

$$v_i = \sum_{g \in G_i^h \cup G_i^a} p_g(1 - p_g)x_g \tag{1.41}$$

$$x \in X, \tag{1.42}$$

where the decision variables, in addition to $x = \{x_g, g \in G\}$, include μ_i and v_i which encode the mean and variance of the number of wins for team i in the shortened season, respectively. Moreover, the following parameters represent the mean and variance of the number of wins for team i in the full season, respectively:

$$\hat{\mu}_i = \mathbb{E}_\xi \left[\hat{y}_i(\xi)\right] = y_i^0 + \sum_{g \in G_i^h} p_g + \sum_{g \in G_i^a} (1 - p_g)$$

$$\hat{v}_i = \mathbb{V}_\xi \left[\hat{y}_i(\xi)\right] = \sum_{g \in G_i^h \cup G_i^a} p_g(1 - p_g).$$

Proof. Proof. See Appendix A.1. □

1.6 Computational Experiments

We performed comprehensive computational experiments to assess the performance of our predictive and prescriptive models. We coded our predictive models in `Python 3.7` using the `scikit-learn` package (Pedregosa et al. 2011). For the prescriptive models, we coded the mathematical models in `C#` and solved the mixed integer programs using the `ILOG Concert` library and `CPLEX 12.10` solver with all solver settings left at their default values. For visualizing the distribution of performance metrics, we used the `vioplot` R package. All experiments were conducted on a Dell desktop equipped with Intel Core i7-6800K at 3.40 GHz CPU and 16 GB of memory running a 64-bit Windows 10 operating system.

1.6.1 Dataset Description

We use historical data from 14 NBA regular seasons (2004–2010, 2012–2018), which are precisely the years that the regular season had the same structure as today; that is, 30 teams, each playing 82 games with schedules constructed using the same formula we described in section 1.1. Prior to 2004, the NBA consisted of fewer teams and as a result the regular season schedule had a different structure. We also omit 2010–11 from our study, since this regular season was suspended for two months, and a 66-game shortened season was adopted.

We used the box score datasets publicly available for all seasons on NBA's official website (NBA 2020c) which contains the detailed information for each game, team and player statistics. We then created 14 training datasets considering three different alternatives for the day at which the season is suspended (i.e., days 100, 120, 140 of the regular season). Note that each regular season takes between 170–180 days. The features input to our predictive model are listed in Table 1.4; see section 1.4.1.

1.6.2 Predictive Model Results

To select a predictive model, we evaluate several binary classifiers using both the widely-used *accuracy* metric and our variant, the *predictive power* metric, as defined in section 1.4.1. We believe accuracy is a good model selection metric when the number of scenarios in the optimization model is small (it measures the predictive accuracy of a single "best" scenario), whereas our predictive power metric is a more powerful model selection metric when the number of scenarios in the optimization model grows large (recall from Proposition 1.1, predictive power may be interpreted as an "expected accuracy" which in the limit is achievable by the prescriptive model when the number of scenarios grows large).

According to Figure 1.6, which shows the distribution of accuracy and predictive power for 5 different classifiers across 14 NBA seasons, while all classifiers exhibit similar accuracy values, the predictive power for the Naïve Bayes classifier is significantly higher than the other classifiers. Hence, we build our prescriptive approach based on the Naïve Bayes predictions. For details of the Naïve Bayes classifier (i.e., predictive model and parameter estimates), see Appendix A.2.

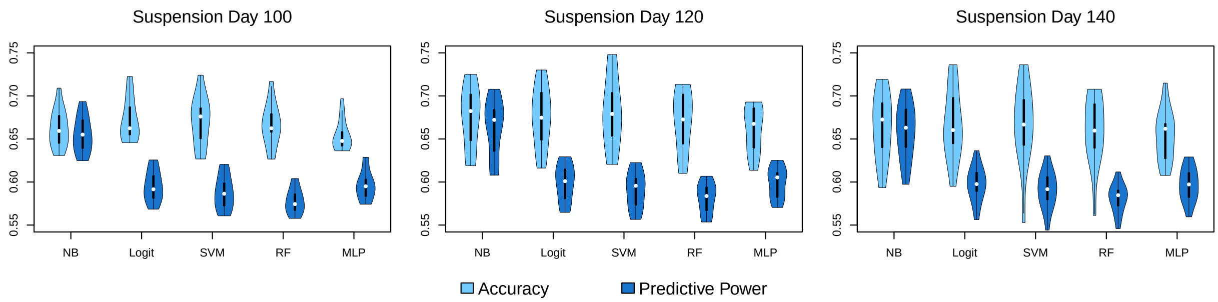


Figure 1.6: Comparison between accuracy and predictive power for 5 different classifiers across 14 NBA seasons

1.6.3 Prescriptive Model Results

In this section, we present the computational results for our proposed prescriptive models PC-MVP, PC-SAA, PM-MVP, PM-SAA, and PW-DQIP. We start by analyzing the impact of the number of scenarios on our SAA-based models. Figure 1.7 presents the performance of PC-SAA across four choices of sample size $|\mathcal{S}| \in \{5, 15, 25, 50\}$. Each boxplot corresponds to 50 replications of the SAA algorithm. The concordance values are obtained after evaluating the solution x proposed by each model on 10,000 randomly-generated scenarios. The panel on the right presents the optimality gaps of the SAA problems after reaching a time limit of 500 seconds. As the number of scenarios increases, one should expect to obtain a closer approximation of the true stochastic problem via SAA. However, a larger sample amounts to solving a more challenging SAA problem. As depicted in Figure 1.7, initially as the sample size increases, the quality of the solution improves in the simulation phase. However, after surpassing 25 scenarios, the SAA becomes computationally intractable, amounting to large optimality gaps and a degradation in the quality of the solution. Hence, the trade-off between the quality of the solution and the runtime is balanced at 25 scenarios. Thus, we select 25 scenarios for the SAA counterparts of our PC and PM models in our main experiments.

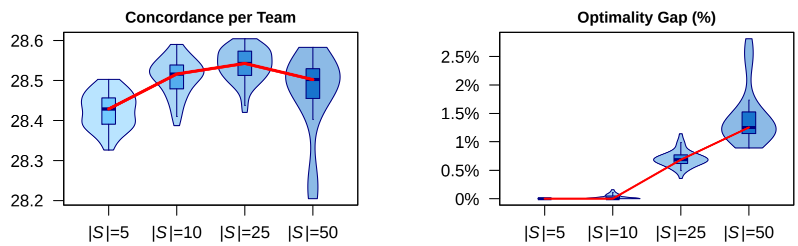


Figure 1.7: Performance of the SAA algorithm across different choices of sample size.

Table 1.5 summarizes the results of applying our variable fixing technique, introduced in Proposition 1.3, in prescriptive models involving the z_{ij} decision variables (i.e., PC and PM). Note that the set of fixed variables is the same between PC and PM models in both solution methods MVP and SAA. The high percentages under the columns “Percentage” in Table 1.5 highlight the effectiveness of the variable fixing technique in eliminating a large proportion of the z-variables across different scenarios in both MVP and SAA. More importantly, the technique is able to eliminate between 7,000–9,000 binary variables in the SAA problems. We also observe that as the suspension day increases (i.e., the season is suspended later), more pairs of teams become impossible to switch ranking positions, given the limited number of remaining games. For instance, when the season is suspended on day 140, the prescriptive model has control over only 17% of these binary variables in a 74-game shortened season, with the remaining 83% of the variables fixed (i.e., eliminated).

Table 1.6 reports the runtime in seconds, objective function value, and optimality gap for all five prescriptive models under different suspension days and target number of games, averaged over 14 seasons. All MVP instances are solved to optimality. For SAA, most instances have not converged to optimality by the 500-second time limit; hence, we report the optimality gaps. For the sake of interpretability of the gaps from our PM-SAA model, inspired by Proposition 1.2, we transform the lower-bound (LB) and upper-bound (UB)

Sus. Day	MVP		SAA	
(Target)	Percentage	Variables	Percentage	Variables
100 (66)	67.4%	293.1	64.5%	7,009.9
100 (70)	72.4%	315.1	70.3%	7,642.9
100 (74)	79.9%	347.4	78.4%	8,527.8
120 (66)	72.5%	315.2	69.6%	7,573.1
120 (70)	74.3%	323.1	72.0%	7,835
120 (74)	80.1%	348.4	78.2%	8,508.9
140 (70)	82.4%	358.3	80.6%	8,761.3
140 (74)	83.9%	365.1	82.3%	8,953

Table 1.5: Number of z-variables eliminated by our variable fixing technique, reported for different suspension days (100, 120, 140) and target number of games/team (66,70,74). Results are averaged over 14 seasons.

obtained by the solver using the formulae $n-1-LB$ and $n-1-UB$, respectively, yielding the optimality gap of $\frac{UB-LB}{n-1-UB}$, which is on the same scale as the gap from PC-SAA. What we notice is that as the suspension day increases, the solution space becomes smaller with fewer games to choose from; thus, all models perform better. Moreover, SAA and PW-DQIP are antagonistic with respect to the target number of games; as the target increases, SAA is able to close the gap more easily, whereas the problem becomes more challenging for PW-DQIP.

Sus. Day	PC-MVP		PC-SAA			PM-MVP		PM-SAA			PW-DQIP		
(Target)	Time	Obj.	Time	Gap	Obj.	Time	Obj.	Time	Gap	Obj.	Time	Gap	Obj.
100 (66)	0.12	29	500	1.43%	28.59	0.10	0	500	2.48%	0.70	500	1.93%	52.9
100 (70)	0.10	29	500	0.88%	28.75	0.08	0	500	1.40%	0.40	500	4.00%	31.3
100 (74)	0.07	29	473	0.21%	28.94	0.08	0	500	0.72%	0.21	500	7.85%	17.6
120 (66)	0.10	29	500	1.19%	28.63	0.09	0	500	2.84%	0.82	147	0.13%	90
120 (70)	0.08	29	500	0.77%	28.77	0.08	0	500	1.61%	0.47	427	1.00%	45.4
120 (74)	0.07	29	491	0.38%	28.88	0.06	0	500	0.95%	0.28	500	3.76%	22.7
140 (70)	0.09	28.9	500	0.86%	28.50	0.07	0.11	500	1.34%	0.61	1.34	0.00%	106.5
140 (74)	0.07	29	500	0.53%	28.76	0.06	0	500	0.90%	0.32	35.5	0.00%	36.7

Table 1.6: Performance of the prescriptive models, averaged over 14 seasons

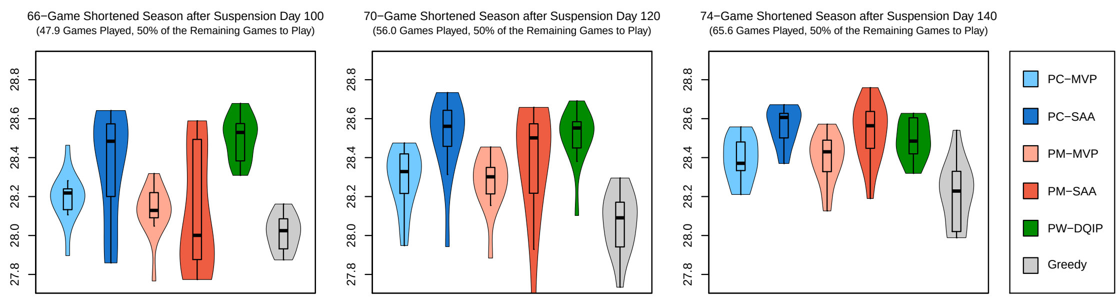
1.6.4 Monte Carlo Simulation Results

Recall that the output of our prescriptive models PC-MVP, PC-SAA, PM-MVP, PM-SAA, and PW-DQIP, is a shortened season x. As a benchmark, we also implement a greedy heuristic which selects games according to their original scheduled dates, with earlier games assigned first until the constraints on the number of games per teams are met. To isolate and measure the performance of the prescriptive model, independent of the predictive model, we assume that the Bernoulli distributions fit by the predictive model are correct. Since we are interested in measuring the expected performance of each model, we obtain a statistical bound (c.f., Kleywegt et al. 2002) on the expected performance metric by conducting a Monte Carlo simulation using a large sample, i.e., we generate 1,000 game outcomes from these Bernoulli distributions. Each realization yields two rankings, one at the end of the shortened season x, and the other at the end of the full season assuming all games are played. We then compare these rankings using our two main metrics, namely, the number of concordant pairs and the Manhattan distance between rankings.

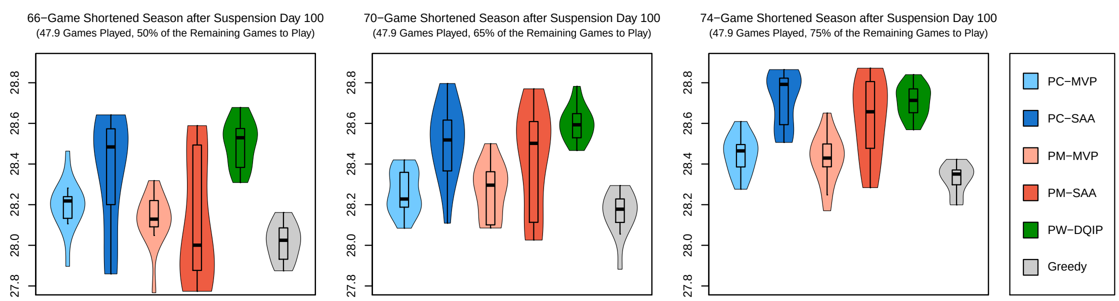
Figure 1.8(a) illustrates the effect of increasing the suspension day, while Figure 1.8(b) shows the impact of using a different target length of the season. Each plot in Figure 1.8 represents the distribution of the simulation results across 14 NBA seasons when measuring the number of concordant pairs per team, and the inner boxes represent the boxplots. Results for the Manhattan distance metric are similar, and so we omit them here but provide them in Figure A.2 in Appendix A.3.1.

According to Figure 1.8, in terms of the average concordance per team, all five proposed models outperform the baseline greedy algorithm in most cases. Comparing the SAA and MVP approximation schemes, we observe that the SAA counterparts of the stochastic models PC and PM perform considerably better than their MVP counterparts, which highlights the effectiveness of SAA in approximating the true underlying distributions. This is particularly pronounced for PC, for which the SAA counterpart dominates MVP in all cases.

For our top three prescriptive models (i.e., PC-SAA, PM-SAA, and PW-DQIP), we perform paired t-tests for each pair of models to quantitatively measure which model outperforms



(a) Target fixed at 50% of the remaining games (i.e., shortened season with 66, 70, and 74 games)



(b) Suspension day fixed at 100

Figure 1.8: Distribution of the simulation results (concordance per team) across 14 NBA seasons

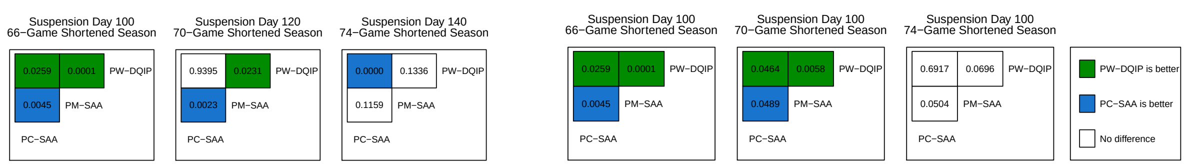
the other. For any two models, m_1 and m_2, our hypothesis test is as follows:

$$H_0 : \mu_{m_1} - \mu_{m_2} = 0 \qquad \text{Null Hypothesis}$$

$$H_a : \mu_{m_1} - \mu_{m_2} \neq 0 \qquad \text{Alternative Hypothesis}$$

The outcome of each test using significance level $\alpha = 5\%$ is one of the following three cases: (i) we cannot reject the null hypothesis, i.e., p-value $\geq \alpha$ which means m_1 and m_2 have statistically similar performance, (ii) we reject the null hypothesis and $\mu_{m_1} - \mu_{m_2} > 0$, i.e., we conclude model m_1 is better, or (iii) we conclude model m_2 is better. Figure 1.9 shows that PC-SAA and PW-DQIP each outperform PM-SAA in three out of five instances, and are tied in the remaining two (note that the leftmost panels are identical in Figures 1.9(a) and 1.9(b)). Moreover, PW-DQIP yields higher concordance than PC-SAA in two instances, is lower in one instance and the two models are tied in the remaining two instances. We observe similar results when comparing our top three models using the Manhattan distance metric; see Figure A.3 in section A.3.1.

To summarize, we observe that PW-DQIP predominantly outperforms PC-SAA, particularly when the league is suspended earlier, i.e., there are more remaining games which makes the instance more computationally challenging for SAA. Recall that while our SAA models approximate the optimization problem using samples but maintain a ranking-based objective (e.g., concordance), PW-DQIP approximates the objective function but explicitly models the full distribution of game outcomes. In our tests, we have found it is generally best to approximate the objective function rather than the distribution of outcomes. Furthermore, PC-SAA tends to produce better solutions than PM-SAA when evaluated on both the Manhattan distance and concordance objectives; this we conjecture is due to the fact that PC-SAA closes the optimality gap faster than PM-SAA.



(a) Instances corresponding to Figure 1.8(a)

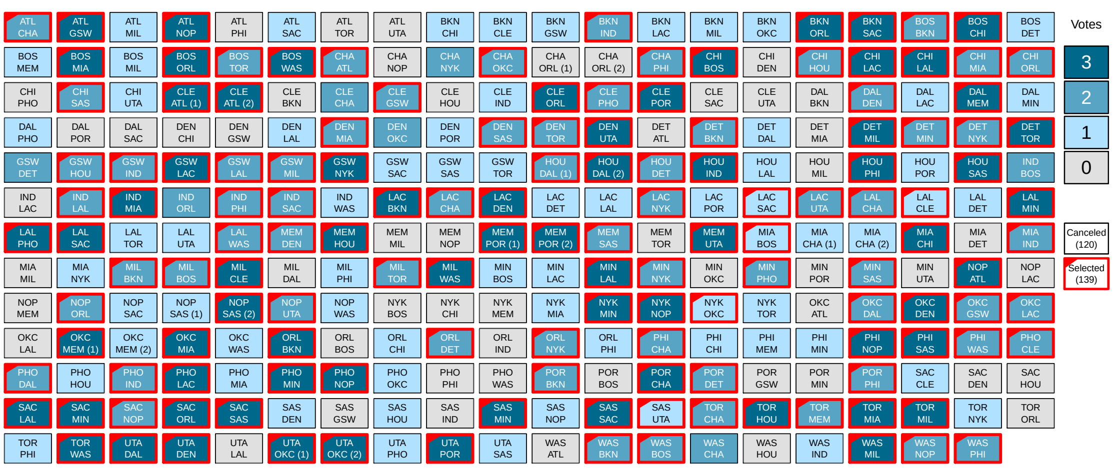
(b) Instances corresponding to Figure 1.8(b)

Figure 1.9: Results of paired t-tests and the p-value for our three top performing models

1.6.5 Suggestions for the 2019–20 Season

In this section, we present the results of our two-phase analytics approach applied to the 2019–20 NBA regular season which was suspended on March 11, 2020. We consider 74 games per team as the target length of the shortened season, thus canceling 8 games per team from the remainder of the season. As a result, out of 259 remaining games, we select 139 games to be played in the shortened season. After the COVID suspension, the NBA was considering multiple alternatives to resume the league, including playing games in a single venue with no fans. The ESPN Wide World of Sports Complex, a Walt Disney property in central Florida, was the front-runner for hosting the remainder of the 2019–20 season (NBA 2020b), and it eventually hosted the NBA shortened season and playoffs in 2020. To model this possibility, in addition to providing results for our base model which has home/away considerations, we also produce an alternative model which replaces the constraints on the number of home/away games per team (1.8) with constraints on the total number of games per team.

(a) Policy 1: 74-game shortened season with home/away considerations



(b) Policy 2: 74-game shortened season without home/away considerations

Figure 1.10: Selected games for the remainder of the 2019–20 season.

To present robust recommendations, we produce an ensemble model which tallies the votes from our three best prescriptive models (i.e., PC-SAA, PM-SAA and PW-DQIP). Figures 1.10(a) and 1.10(b) show the number of models which select each game when home/away considerations are in force (policy 1) and relaxed (policy 2), respectively. For each policy, we solve an optimization problem using the number of votes for each game as weights to obtain a feasible game plan while maximizing the total number of votes. The games with red borders are our final selections. We find that 101 games are selected by both policies and 82 games are selected by neither policy.

To analyze the projected ranking at the end of the shortened season, we calculate the expected number of wins by each team using the final game plans shown in Figure 1.10, and obtain the projected rankings for each policy based on the expected number of wins. The results, as shown in Figure 1.11, not only confirm the effectiveness of our predictive model, but also address some of the fairness challenges regarding the conclusion of the current season widely discussed by the media and mentioned in section 1.3.1. We compare our projected rankings in Figure 1.11 with the ranking at the time of suspension from Figure 1.3 in section 1.3.1. Doing so, we find that in both of our projected rankings, the `Pelicans` replace the `Grizzlies` as the last (8^{th}-ranked) team to qualify for the playoffs in the western conference. The improved performance of the `Pelicans` in the remainder of the season is widely anticipated by pundits, owing to the fact that the `Pelicans` started to improve their performance during the weeks leading to the suspension after their number one pick rookie, *Zion Williamson*, rejoined the team after missing the first 50 games due to injury. The toughness of the remaining schedule is also reflected in our results. The `Grizzlies` and the `Nuggets`, who both have a difficult remaining schedule, are ranked lower in both our projected rankings, as compared to their ranking at the time of suspension.

Our analysis also shows that there can be significant changes to the end-of-season ranking if the league chooses, for logistical reasons, to hold all games in one physical location. As an example, the `Philadelphia 76ers`, known to be a dominant team at home with a home

Western Conference	League (West)		Eastern Conference	League (East)	
	Policy 1	Policy 2		Policy 1	Policy 2
Los Angeles Lakers	**1** (1)	**2** (1)	Milwaukee Bucks	**2** (1)	**1** (1)
Los Angeles Clippers	**3** (2)	**3** (2)	Toronto Raptors	**4** (2)	**4** (2)
Denver Nuggets	**13** (7)	**13** (7)	Boston Celtics	**5** (3)	**5** (3)
Utah Jazz	**9** (4)	**9** (4)	Miami Heat	**7** (5)	**6** (4)
Oklahoma City Thunder	**8** (3)	**7** (3)	Indiana Pacers	**11** (6)	**12** (6)
Houston Rockets	**12** (6)	**10** (5)	Philadelphia 76ers	**6** (4)	**8** (5)
Dallas Mavericks	**10** (5)	**11** (6)	Brooklyn Nets	**16** (8)	**16** (8)
Memphis Grizzlies	**20** (12)	**17** (9)	Orlando Magic	**15** (7)	**15** (7)
Portland Trail Blazers	**18** (10)	**18** (10)	Washington Wizards	**25** (11)	**23** (10)
New Orleans Pelicans	**14** (8)	**14** (8)	Charlotte Hornets	**26** (12)	**25** (12)
Sacramento Kings	**19** (11)	**20** (12)	Chicago Bulls	**28** (14)	**28** (14)
San Antonio Spurs	**17** (9)	**19** (11)	New York Knicks	**22** (9)	**22** (9)
Phoenix Suns	**21** (13)	**21** (13)	Detroit Pistons	**29** (15)	**29** (15)
Minnesota Timberwolves	**24** (14)	**26** (14)	Atlanta Hawks	**23** (10)	**24** (11)
Golden State Warriors	**30** (15)	**30** (15)	Cleveland Cavaliers	**27** (13)	**27** (13)

Figure 1.11: Projected rankings based on policies in Figure 1.10 (conference ranks are inside parentheses)

win ratio of 29/31 and road win ratio of 10/34, is ranked 6$^{\text{th}}$ in the league by policy 1 (4$^{\text{th}}$ in the East), while ranked 8$^{\text{th}}$ in the league by policy 2 (5$^{\text{th}}$ in the East). In policy 2, the `76ers` will lose home court advantage in the playoffs, as opposed to policy 1 which grants them the right of hosting first in the playoffs.

1.7 Discussion, Limitations and Future Work

On occasion, world events such as COVID-19 make shortening a sports season necessary. When this occurs, it is difficult if not impossible to select a subset of games which all parties – the league, the teams, and the fans – will consider a "fair" compromise. While our model is not perfect, we believe that by maximizing the expected concordance between the shortened season's ranking and that of the full season, we may produce a shortened season which is as similar to the full season as possible, and by this measure may be deemed as the fairest practical compromise.

Moreover, it is important to note that our methodology is very different from one which takes the ranking at the suspension date and tries to sustain this ranking through the end of the shortened season. Despite the fact that our predictive model uses pre-suspension data to calibrate our model's parameters, and our prescriptive model optimizes based on these parameters, it is also the case that a team with a high win rate pre-suspension will not necessarily have a high win rate post-suspension. The nature of the shortened season and its impact on rankings is much more complex, owing to the relative difficulty of the schedules before and after the suspension (i.e., precisely when a team faces easy or hard-to-beat competitors). Indeed, because our prescriptive model aims to produce a shortened season with a ranking similar to that of the full season, if a team had a relatively easy schedule pre-suspension it would generally have a relatively difficult schedule post-suspension, and our prescriptive model will naturally attempt to maintain this difficult schedule in the shortened season post-suspension. Figure 1.12 shows that the ranking at the time of suspension is markedly different than the rankings of our shortened seasons, and moreover, concordance (relative to the end-of-season ranking) is much higher for our shortened seasons.

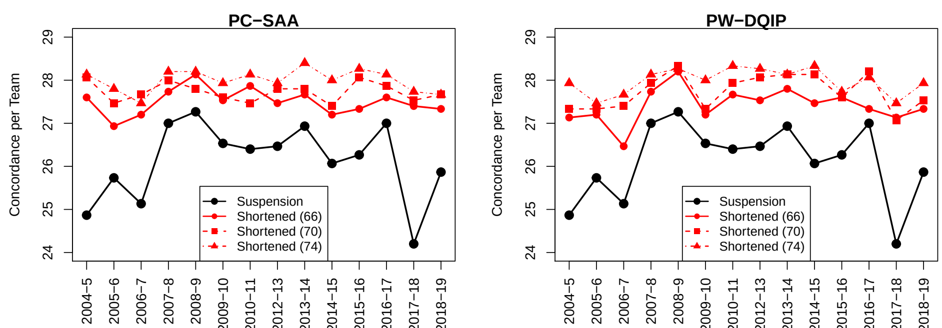


Figure 1.12: Comparison between the ranking at the time of suspension and the shortened season rankings (at various target number of games/team). Concordance of all rankings listed is with respect to the end-of-season ranking. Suspension day is 100 with 48 games on average played per team pre-suspension.

We envision several directions for the future research. First, under different circumstances, apart from the fairness of the schedule, other objectives (e.g., travel cost and distances, broadcasting restrictions, venue availability) may be deemed as the top priority. Alternatively, the problem can be framed as a multi-criteria decision making problem. Second, more sophisticated predictive models may help in devising more informative decisions. For instance, one can exploit more accurate representative features (e.g., player-level statistics) or temporal aspects (e.g., chronological dependencies between the games) in the predictive analytics phase. Moreover, when a league suspension occurs before the season starts (as is the case with most player strikes), historical data from the previous season(s) would be required to fit the predictive model's parameters. Third, from a computational point of view, solution methods based on large-scale optimization techniques (e.g., generalized Benders decomposition) may be designed to tackle large instances of the stochastic optimization problems (e.g., SAA with a larger sample).

1.8 Conclusion

Professional sports leagues may be suspended due to various reasons, requiring the league to select which games to play in a shortened season. In this study, we proposed a two-phase analytics approach for this problem. In phase one, we predicted game outcomes using a Naïve Bayes classifier. In phase two, we used stochastic optimization techniques to prescribe a data-driven decision which maximizes the expected similarity between the ranking at the end of the shortened season and the full season had it been played in full. To solve two of our stochastic optimization problems (PC and PM), we proposed approximation schemes (MVP and SAA), and variable fixing techniques. For our third model (PW), we introduced a deterministic equivalent reformulation (i.e., PW-DQIP). Our SAA models approximate the distribution but have an exact objective, while PW-DQIP has an exact distribution but approximates the objective. We evaluated the quality of the solutions produced by each approach using Monte Carlo simulation. Our computational experiments suggest that PW-DQIP outperforms SAA even for reasonably-large 25-scenario instances. Finally, we suggest two alternative policies for the remainder of the 2019–20 NBA season which differ in whether teams play their games in the original cities as-planned or are played all at the same venue.

Chapter 2

Analysis of Pricing Mechanisms in a Resource Exchange Economy

2.1 Introduction

In recent years, resource sharing has become increasingly popular in multi–agent settings such as airline alliances, liner shipping alliances, and equipment sharing platforms (e.g., construction, health care, and scientific laboratories). With the rapid growth in size of such industries in the past 100 years, forming alliances have become more popular, and alliance members can never achieve such increased profitability by focusing on their internal decisions alone. According to Hu et al. (2013), around 75% of all the passengers in the world have flown with one of the three major airline alliances (i.e., Star, Sky Team, and Oneworld). According to Hingorani et al. (2005), the top 10 sea cargo carriers control about 80% of the sea transportation market. In the health care sector, medical equipment sharing has helped hospitals and health service providers save millions of dollars annually (Sanborn 2018, Cohealo 2021). Prior to the emergence of novel resource sharing platforms, companies had been either purchasing and taking the ownership of their resources and equipment, or renting equipment from rental companies (e.g., Herc Rentals, Yard Club, United Rentals in heavy equipment rentals) for specific uses and periods of time.

There are many reasons why companies would like to share their resources as it helps each member of a formed alliance among these companies to reduce overall cost, lower lead time, and increase asset utilization. All of these benefits become crucial for profitability especially given the challenges of owning equipment or renting equipment from rental companies. With the significant growth in use of recent technological advancements, tools and equipment are becoming heavier, bulkier, and more expensive, while also delivering better quality services and multiple tasks. Therefore, nowadays owning an equipment comes with major risks and possible costs. *Affordability* is one of these challenges. Whether business owners can afford purchasing *specialty equipment*, which are usually very expensive, is oftentimes a very complex question. A potential owner of a specialty equipment needs to assess the usage frequency and the potential benefits that can be generated and decide whether the benefits match the purchase cost. Each company, at some point, needs a specialty equipment for a specific job, and if she owns this specialty equipment, it sits idle for months or even years after the job is complete. Of course she needs to take into account depreciation cost, and the fact that not only idle equipment does not contribute to revenue of the company, but also adds to *maintenance* and inventory holding costs. Renting equipment from rental companies also presents its own challenges. The main challenges are the lack of control over *equipment availability* and potential *high volatility in prices.* Another problem with rental companies is their sensitivity to economic conditions which might leave them as unreliable options at the times of economic difficulties. After the COVID–19 pandemic for instance, and almost a year of inactivity in some businesses, several rental companies have filed for bankruptcy (Frommer's 2021).

To counter these risks, the idea of *sharing economy marketplaces* can be appealing. Instead of limiting ourselves to a strict dichotomy (e.g., to purchase or to rent from rental companies), there is a third option which is to exchange resources between businesses who are at the same level of hierarchy in the economic system. By borrowing necessary equipment from other businesses (e.g., construction companies, medical facilities), and paying for the actual usage time is much more affordable compared to ownership. Maintenance is less of an issue either, since the maintenance cost can be shared among all businesses that use a particular equipment. Whether the resource to be shared is consumable and/or perishable (e.g., a seat on a flight, a segment of a container on a sea cargo carrier, wood or metal for a manufacturing project) or a specialty equipment (e.g., an expensive medical equipment) which is shared by other agents for an allocated time frame, we can always quantify the amount of such shared resources. As an example, one unit of a shared resource can be 1 pound of a fluid material, 1 wooden desk, or 1 hour of usage time on an equipment. In the next section, we will discuss how resource sharing decisions are made and can be coordinated in practice.

2.1.1 Resource Pricing Problem

In an economic system where there are multiple agents (e.g., airlines, liner shipping companies, equipment owners) endowed with multiple types of resources (e.g., seats on flights, specialty equipment), resource sharing decisions are typically made in a decentralized way, i.e., each agent make its own decision independent of or even oblivious to those of others'. A major approach to coordinate these decisions is through resource pricing mechanisms. The *pricing mechanism*, which is the announcement of a resource price function by a central planner, followed by a settlement process, is the means by which decisions of agents and the central planner interact to determine the distribution of resource endowments among agents. To design a pricing mechanism, a key question is *What resource exchange price function will induce decentralized agents to exchange their resources in such a way as to maximize the sum total of the profits generated by all agents?* In other words, *under what resource pricing* *policy will the mismatch between supply and demand due to the conflict of interest among decentralized agents be at a minimum?*

Specifically since the decentralized problem is a collection of individual optimization problems, when combining the production and exchange quantities, the total supply and demand of different resources may or may not match. Moreover, decentralized resource exchange decisions reflect individual profit maximizing incentives, which may not necessarily give rise to efficient resource allocations for the entire system. To address these problems, resource prices would be chosen in such a way that the collection of agents' solutions match the centralized solution (also known as the *first best solution*), in which case the solution of the decentralized problem is called a *coordinating solution*. In case of any mismatch, unsatisfied buy/sell orders can be addressed using spot market options as it is often the case in practice. We call this process the *settlement process*. This helps satisfy all agents' resource needs and restores the feasibility of the resource exchange solution. Note that in this case, due to additional transaction and prices paid to the spot market, the total profit may be lower than that under the first best solution. The decentralized solution in this case is called an *approximate coordinating solution*. Figure 2.1 shows the sequence of events by the central planner and the agents (i.e., those businesses involved in the alliance and would like to exchange resources) that we consider in this chapter.

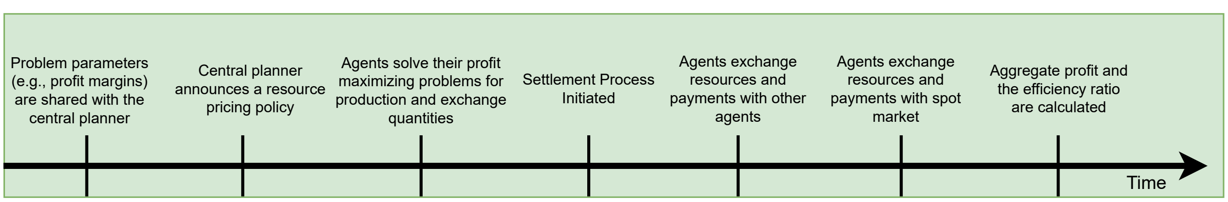


Figure 2.1: Timeline of events in the coordination problem

There are two main streams of papers studying the effectiveness of resource pricing policies. Some studies consider the best–case decentralized solution in comparison to the first best solution (Hu et al. 2013, Agarwal and Ergun 2008, Chun et al. 2016). Other papers study the worst–case decentralized solution (Baumol and Fabian 1964, Jennergren 1972, 1973, Houghtalen et al. 2011, Roels and Tang 2016). We focus on the latter case and take into account the possibility of supply-demand mismatch in decentralized resource sharing, which leads to approximate coordinating solution for the decentralized problem in the worst case.

Our resource price function may be linear or nonlinear. In a linear price function, a simple price value for each resource is announced which allows agents to buy or sell any amount of those resources at that price tag. Any other form of a price function with a price schedule, varying by the quantity bought or sold, is called nonlinear pricing (Van Zandt 2012). There are advantages and disadvantages to nonlinear pricing. With a nonlinear price function, we can ensure that all agents have unique optimal solutions, thus helping to coordinate the agents to a globally optimal solution. On the other hand, a nonlinear price function is less intuitive and more difficult to communicate, compared to a linear price function. Specifically, as we will show in later sections, with a nonlinear price function, all agents' problems are guaranteed to produce a unique optimal solution. Hence, nonlinear price function provide a method of controlling the outcome of the decentralized problem. Apart from nonlinearity of the resource price function, *price discrimination* is another way of controlling the agents' decisions. Price discrimination exists when identical resources are bought and sold by the agents at different prices. For the worst–case decentralized solution, it is well–established that linear pricing without price discrimination cannot coordinate the agents to produce exactly the first best solution (Baumol and Fabian 1964). According to our results, with price discrimination, linear pricing can always achieve an efficiency ratio of one, circumventing the multiple optimal solutions problem in agents' problems by choosing different prices for different agents. On the other hand, Jennergren (1972) has proved that a quadratic price function with price discrimination can achieve the first best solution in the decentralized problem. Our goal in this chapter is to study whether a price function without price discrimination can be designed to achieve the first best solution in the decentralized problem. Figure 2.2 illustrates the four cases mentioned with the focus of this chapter highlighted in red in the bottom right corner.

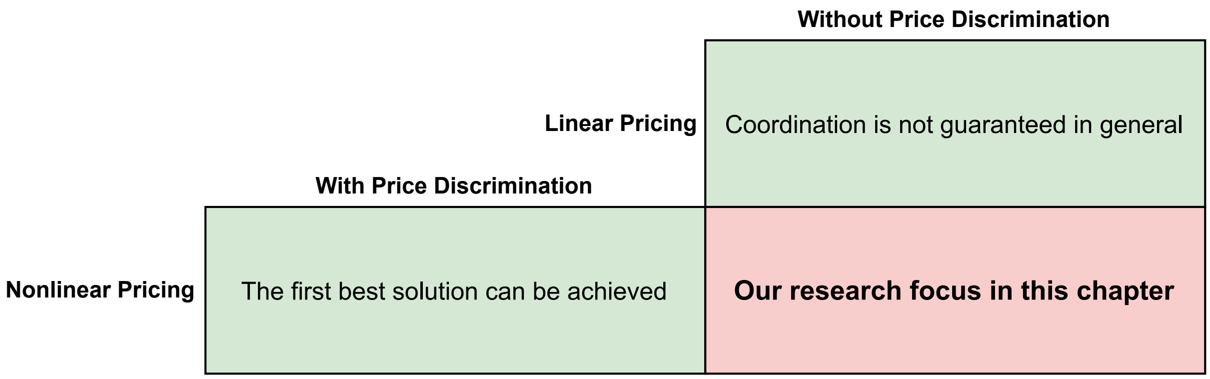


Figure 2.2: Four possible cases with respect to the shape of the price function and the existence of price discrimination

Our main contributions to the literature of operations research and resource sharing is two–fold:

1. We establish the effectiveness of nonlinear pricing mechanisms with no price discrimination. Specifically, we show that the first best solution can be achieved in the decentralized system under such pricing mechanisms under certain conditions, and we provide a full characterization of those conditions.

2. We highlight the usefulness of nonlinear price functions in improving the system efficiency of decentralized resource exchange decisions under uncertainty.

The remainder of this chapter is organized as follows. We begin with an overview of literature in section 2.2. In section 2.3, we introduce the mathematical formulations for the centralized, decentralized and the overall pricing mechanism problem, and we present the main properties of each model. We introduce the concept of efficiency ratio as a metric to gauge the effectiveness of any price function. We analyze the efficiency ratio that can be achieved under quadratic price functions without price discrimination. In section 2.5, we introduce an extension of the model in which the profit margins of the agents are not fully–known and the central planner only knows a probability distribution of those coefficients while designing the resource prices. We present our numerical analysis in section 2.4. We conclude this chapter in section 2.6.

2.2 Literature Review

The Oxford English Dictionary[1] has defined the *Sharing Economy* as *an economic system in which assets or services are shared between private individuals, either free or for a fee, typically by means of the internet.* Since our focus in this study is on those instances in which a payment is required for sharing resources, we use the term *resource exchange* to distinguish between the two cases (i.e., sharing resources for free and for a fee). In other words, resource exchange is the practice of sharing resources, owned by (or endowed to) players of an economic system, when the amount of exchange is determined by a cost/benefit analysis, which typically involves solving an optimization problem, and it depends on the resource price function. Players will then use the resources to produce a product or a service package to sell to the outside market. Therefore, there is one source of cost (i.e., buying additional resources from other players) and two sources of profit (i.e., selling endowed resources to other players and selling the final product to the customers). In this section, we review the extensive literature on the coordination problem in resource exchange models with different configurations, assumptions, methodologies and results.

Owen (1975) has defined the linear production game in a cooperative game setting. Our problem, even though similar in spirit to the linear production game, uses non–cooperative game theory to address the coordination problem. The resource exchange setting with linear production (i.e., linear resource constraints and linear profit function) is studied in detail in Baumol and Fabian (1964) and Jennergren (1972). The main players are: i) the central planner (e.g., alliance manager in airline or liner shipping alliances, platform designer for

[1] See `https://www.lexico.com/definition/sharing_economy`

equipment sharing), and ii) agents, as economic entities, endowed with certain amount of resources, each having a different profit coefficients and resource usage rates. Much of the attention in the operations research community has gone to the case where multiple agents buy resources from a central pool, including in applications such as the *unit commitment problem* in electrical power production, and the *federated cloud computing problem* in traffic management of cloud computing systems (Feizollahi et al. 2015, Fu et al. 2005, Maheswari and Vijayalakshmi 2012, Gomes et al. 2012, Voos 2007, Johari and Tsitsiklis 2003, Maillé and Tuffin 2008, Cachon 2003), and not enough has been discussed about exchanging resources directly between agents (Guo et al. 2007, Houghtalen et al. 2011, Hu et al. 2013, Chun et al. 2016, Roels and Tang 2016). We use the term *resource exchange* to distinguish between exchanging resources across agents in one hand, and the traditional model of buying from a central pool. In the remainder of this section, we review the main streams of papers in the literature.

Game theory and contract design concepts have been used extensively in the resource sharing and resource exchange literature to address the coordination problem in multi–agent economic models. Roels and Tang (2016) investigate bidirectional alliances that occur between only two companies. They discuss a bargaining framework to allocate resources and the total revenue between a manufacturer and a marketer in the car manufacturing sector. Houghtalen et al. (2011) introduce an inverse optimization method to address the coordination problem using duality theory, and compare their solution with the Nash bargaining solution. While employing non–cooperative game theory techniques, they make two key assumptions about the behavior of consumers, each resulting in a different analytical solution for the decentralized problem. One method of controlling the behavior of the agents, which is relatively common in the literature, is to consider an upper bound on the amount of resource exchanges each agent can make. Chun et al. (2016) study alliances consisting of competing firms. Their main research question is whether it is guaranteed that all alliance members always benefit by being a member of the alliance, and if so, how the profit should be allocated among agents. They propose an alliance structure with two main properties: i) it allocates resources among agents using a Mathematical Program with Equilibrium Constraints (MPEC), and ii) their allocation satisfies several axioms including *Pareto optimality*. Perakis and Roels (2007) study a contract form called the *price–only contract* (i.e., a linear price function), and examine its efficiency in different supply chain configurations (e.g., push or pull inventory positioning). Hu et al. (2013) introduce a two–stage analysis for the problem of sharing revenues in airline alliances. In stage one of the analysis, airlines as agents negotiate the terms of the agreement and revenue sharing rules. In the second stage (operational level), each airline solves her own optimization problem, and announces her own resource sharing policy. Revenue is distributed according to the terms of the agreement. Agarwal and Ergun (2008) study a multicommodity flow network with multiple players, each with certain amount of capacity on the edges of the network as resource endowment. Players exchange capacities with other players in an alliance. Inverse optimization and a cooperative game theory framework have been used to solve the coordinating problem. Instead of working with detailed contract designs and behavioral assumptions for the agents, our goal is to focus on the simplest contract possible, i.e., a resource price function, which will allow the agents to make production/exchange decisions independently (not sharing the decision making phase with other agents or the central planner) and obliviously (not being concerned about other agents' decisions). This allows us to derive structural results that characterize the advantages of nonlinear vs linear pricing functions and the impact of price discrimination, which has not been studied in the literature before.

Early optimization studies and especially papers addressing the **decomposition of large mathematical programs** have presented great potential to advance the theoretical foundations of the mechanism design problem in resource sharing and resource exchange economic settings. The idea of decomposing linear programs into smaller problems was first introduced by Dantzig and Wolfe (1961). A large number of scholarly works followed suit to advance the theory on decomposition techniques. Baumol and Fabian (1964) proved that with the presence of local constraints, linear pricing cannot solve the coordination problem. Jennergren (1972) proposed the idea of using a nonlinear price function to coordinate such systems. After almost three decades, with the advent of computers, a new attention has been paid to the old decomposition theories, and among those Jose et al. (1997) revisited Jennergren's methods, proving that the idea works even when the agents are assumed to have a nonlinear profit function. Guo et al. (2007) introduce a market–based optimization algorithm, inspired by decomposition algorithms, for the coordination problem in the resource exchange setting. In their analysis, the central planner is called *dealer*, and she owns some endowments, as well. Through an iterative approach, similar to Dantzig–Wolfe decomposition, each agent determines the bundle of resources to trade with other agents in a problem called *sub–problem*, and then the dealer solves the *settlement problem* to resolve the mismatch between supply and demand of different resources and announces new set of resource prices, identical for all agents. This process continues until agents reach the globally optimal solution in a finite number of iterations. In this chapter, even though our proposed price function is inspired by the nonlinear price functions of the form introduced in Jennergren (1972) and Jennergren (1973), we focus on those price functions without price discrimination, and we aim to propose a price function that solves the coordinating problem in one step, without relying on multiple stages of an iterative approach, as in Guo et al. (2007).

Auction theory is another relevant concept that has been used to address the problem of finding the best resource prices. Instead of solving a single–shot optimization problem, auction theory proposes an iterative approach in which the prices as well as resource allocation are being updated until a convergence has been declared (i.e., agents' solutions match the first best, i.e., central planner's, solution). The problem of buying and selling resources has a long history in the chain of great economists. Leon Walras has studied the iterative approach by which sellers and buyers reach an equilibrium state, and he calls the process *tâtonnement* (French for "trial and error") (Walras 1969). Tâtonnement is not guaranteed to solve the coordinating problem between buyers and sellers in a finite number of iterations. Arrow (1951) and Debreu (1951) introduced the concept of *competitive equilibrium* (otherwise known as *Walrasian equilibrium* to acknowledge the works of Leon Walras) which is an equilibrium state in a commodity market with multiple traders and flexible prices. The crucial assumption about this competitive environment is the fact that small exchanges by the agents do not impact the overall resource price function. There are various ways of computing Walrasian equilibrium prices and equilibrium distribution of the resources, and it is a difficult problem in general, which has been studied in economics and computer science communities. One of the prominent methods of finding the prices and resource distribution is through an ascending auction, where the prices are initially set to zero, and they rise over time until the convergence has been declared. This mechanism has few names in auction theory literature including *Walrasian auction* and *English auction*.

Walrasian equilibrium prices have a remarkable property in that they allow each buyer to purchase a bundle of goods that she finds the most desirable, while guaranteeing that the induced allocation over all buyers will globally maximize social welfare. However, there are two caveats. First, the prices may induce indifferences which result in multiple optimal solutions. In fact, the minimal equilibrium prices necessarily induce indifferences. Accordingly, buyers may need to coordinate with one another to arrive at a socially optimal outcome. Indeed, the prices alone are not sufficient to coordinate the market (Arrow 1974). Second, although tâtonnement–type procedures converge to Walrasian equilibrium prices on a fixed population, in practice, buyers typically observe prices without participating in a price computation process (Arrow and Hurwicz 1958). These prices may not be the perfect Walrasian equilibrium prices, but instead somehow reflect distributional information about the market (Arrow and Debreu 1954). Hsu et al. (2016) investigate different conditions under which Walrasian equilibrium prices will be the resulting solution for an auction. They assume specific conditions on the value function used by each agent in their optimization problem, and prove that minimal Walrasian equilibrium prices are optimal under these conditions.

They also study how the central planner learns about agents' problem parameters and their solution dynamics over time and the minimum number of iterations of the auction that is required for the central planner in order to approximately learn the true behavior of each agent in terms of resource demand and production plan. This minimum number is called *sample complexity* which is also a popular topic among computer scientists. Also related to the behavioral assumptions about agents, a mechanism is called *incentive-compatible (IC)* if every player can achieve their best possible outcome just by acting according to her true preferences. Note that there are two types of incentive compatibility: dominant-strategy incentive-compatibility (DSIC) and Bayesian-Nash incentive-compatibility (BNIC). In the first case, one fares best or at least not worse by being truthful, oblivious to others' strategies. According to the second concept however, "if" all other agents are being truthful, the best strategy is to be truthful. It is important to note that the problem setting that we are interested in is of DSIC type in which strategic considerations cannot help achieve a better profit. Malakhov and Vohra (2009) have studied an example of BNIC auction which is explained by a simple network flow. See Ertogral and Wu (2000), for more details.

2.3 Mathematical Model Description

2.3.1 Centralized vs. Decentralized Decision Making

The problem of determining resource exchange quantities can be studied from two different angles:

1. **Centralized Problem**: We can look at the resource exchange problem from a *central planner*'s perspective which considers the combined profits of all the agents as the objective function. In this problem, resource endowments are shared in a single pool, and there is no cost associated with exchanging resources among agents in the objective

function. Therefore, the central planner's problem determines production (and possibly exchange) quantities while maximizing the overall production profit, subject to a set of linear resource constraints.

2. **Decentralized Problem**: We can also look at this problem from each agent's perspective, considering only the net gain of that specific agent as an objective function for planning purposes. Each agent has two components in her objective function:

$$\text{objective function} = \text{production profit} - \text{net payment for exchanged resources}$$

   Agents are assumed to be making their production/exchange decisions, independently of (or oblivious to) other agents' decisions, and this is the main reason why centralized and decentralized solutions might not be in agreement. As an example, for a resource that is in high demand, multiple agents might place buying orders and there may not be enough supply to satisfy all such orders. On the other hand, for a resource with high exchange price, there might be multiple sellers and not enough demand to satisfy the selling requests. In these situations, we assume a *Spot Market* where agents can buy/sell resources, incurring some additional cost. In other words, exchange prices within the economic system are designed in such a way that, it is always beneficial to satisfy buy/sell orders within the network of agents, if possible, rather than relying on the spot market.

We explain the practical relevance of the modeling framework based on the example of airline alliances. Imagine different airlines (e.g., KLM, Delta, American, Korean) are agents, alliances (e.g., Star, Sky Team, Oneworld) are resource sharing platforms or central planners, seats on different flights are resources, and finally flight itineraries are products or services which may require multiple resources. To further illustrate this example, let us go over two of the common resource sharing practices in airline alliances:

- **Codesharing**: A codeshare agreement is a business arrangement in which two or more airlines publish and market a flight under their own airline designator and flight number (the "airline flight code") as part of their published schedule. Typically, a flight is operated by one airline (technically called an "administrating carrier") while seats are sold for the flight by all cooperating airlines using their own designator and flight number.

- **Interlining**: Interlining is an agreement between individual airlines to handle passengers traveling on itineraries that require multiple flight legs on multiple airlines. Such agreements allow passengers to change from one flight on one airline to another flight on another airline without having to gather their bags or check-in again.

Multi-city flight itineraries for passengers and a pack of codeshare and interline seats for travel agencies can be considered a profitable product of an airline in which multiple types of resources are required to be able to offer the final product. That is why the nature of resource sharing in the airline industry is very similar to a production setting where multiple agents exchange different resources to produce their final product, while considering the net profit through the exchange of resources.

In the rest of this section, first we introduce the notation and the basic definitions of the mathematical models. Then, we formulate the central planner's problem (i.e., centralized problem) and agents' problems (i.e., decentralized problem). We discuss possible resource pricing strategies and formulate the problem with a specific form of a nonlinear price function, and we introduce theorems and lemmas characterizing the instances in which a nonlinear price function solves the coordinating solution without price discrimination.

2.3.2 Definitions and Problem Statement

Basic notation: Table 2.1 contains the description of all the basic parameters of the mathematical model (i.e., agents, resources, profit and resource usage coefficients, endowments).

Parameter	Description
J	Set of all the agents
R	Set of all the resources
π_j	Per unit production profit for agent j
a_{ij}	Amount of resource i needed for one unit of product j ($a_{ij} > 0, \forall(i,j)$)
b_{ij}	Endowment amount of resource i for agent j
$\bar{b}_{ij} = \sum_{j' \neq j} b_{ij'}$	Total endowment of resource i owned by all agents except agent j
$b_i = \sum_j b_{ij}$	Total endowment of resource i among all agents

Table 2.1: Parameters of the resource exchange mathematical model

Decision variables: There are two sets of decision variables in this problem: production and resource exchange quantities. The production quantity is denoted by x, and the exchange quantity is denoted by t. We use a superscript c on those variables associated with the central planner's problem.

- x_j, x_j^c = the amount of final product produced by agent j;
- t_{ij}, t_{ij}^c = the net amount of resource i that agent j buys from other agents. A negative value means agent j sells that amount to other agents. Note that this exchange variable is unrestricted in sign, and it is possible to use two non–negative decision variables (e.g., θ_{ij}, γ_{ij}), instead (i.e., $t_{ij} = \theta_{ij} - \gamma_{ij}$, where θ and γ are buying and selling quantities, respectively).

Optimal value of each decision variable are indicated with an asterisk, e.g., x_j^*, t_{ij}^*.

Central planner's objective function: The central planner's objective function value is denoted by z_c and it is equal to the total profit that agents make through production. From the central planner's perspective, the contribution of resource exchanges to the objective function is zero, as the buy/sell prices cancel out in the central planner's problem. The mathematical formulation of the objective function in the central planner's problem is given in (2.1), below.

$$z_c = \sum_j \pi_j x_j^c \tag{2.1}$$

Central planner's resource constraint: There are two ways to formulate the resource constraint in the central planner's problem. The most natural way is to assume all resource endowments of each type to be in the same pool (i.e., amount $b_i = \sum_j b_{ij}$ for each resource i) and distribute that sum among all agents, based on the production quantities.

$$\sum_j a_{ij} x_j^c \leq b_i \qquad \forall i \tag{2.2}$$

According to inequality (2.2), the total amount of resource i used for production purposes cannot exceed the total endowment of resource i.

We can also model the resource constraint, while keeping track of the exchange quantities among agents. After committing exchange decisions, the original endowment values (i.e., b_{ij}) may either increase, decrease, or stay the same, and the new upper bound for the resource usage will be an adjusted bound using the exchange quantities. Constraint (2.3) ensures that each agent can use resources for production purposes up to that adjusted endowment value (i.e., $b_{ij} + t_{ij}^c$). Finally, constraint (2.4) is the *market clearing constraint*, as it makes sure that the total supply and demand of all the resources match.

$$a_{ij} x_j^c \leq b_{ij} + t_{ij}^c \qquad \forall i, j \tag{2.3}$$

$$\sum_j t_{ij}^c = 0 \qquad \forall i \tag{2.4}$$

Linear dual prices: One of the prominent choices in the literature for resource prices in coordination problems is the optimal value of the dual variable for the resource constraint. Since the optimization problem is a linear programming (LP) problem, we call the use of optimal dual values as resource prices, *linear dual pricing*. The dual variables for constraints (2.2), (2.3), (2.4) are p_i, q_{ij}, h_i, respectively.

General form of resource price function: Price function $f_{ij}(x, t, \Phi)$ is a general form of resource price function which depending on our choice of the function, may depend on production and/or exchange quantities, as well as some external parameters denoted by Φ. According to the original scholarly works discussing the general concept of a *price mechanism*, Coase (1937), Saari and Simon (1978), there are mainly three properties for a good price function:

- **Signalling**: changes in the price function, the value and the form, impacts agents' decision making process;
- **Incentives**: the total demand of a resource by all agents may change the price of that resource, which in turn has an effect on the production quantity of all agents;
- **Rationing**: the price function serves as a tool to ration scarce resources when demand is larger than supply.

It is not difficult to see that a quadratic pricing policy, as a function of exchange quantity, satisfies signalling and rationing properties. Regarding the incentive property, note that our goal is to eventually propose a price function inspired by the first best solution, and the desired production quantity in the centralized solution provides enough incentives as demand point for the agents to make decisions in a globally optimal way. Therefore, the quadratic pricing policy is the basis of our analysis in this chapter. We can ensure that with a nonlinear objective function, agents' problems will not have multiple optimal solutions.

Moreover, using *inverse optimization* introduced by Ahuja and Orlin (2001), we can direct all agents' problems to a globally optimal solution using a nonlinear price function. Even though there is no demand constraint in our problem setting, the production acts as a demand point, as the agents may need to acquire more resources to produce. Therefore, a production quantity dictates how much resource exchange needs to happen to facilitate the desired production quantity, and as we will see in the following sections, the optimal exchange quantity in the central planner's problem may impact our choice of the nonlinear price function. Finally, the nonlinear price function of the form described below is the most natural choice to manage any rationing requirements in the case of scarce resources, as the overall payment depends on the amount of exchanged resource.

In the price function (2.5), which is called an *"augmented price function"*, the parameter Φ includes the constant part of the augmented price function, r_{ij}, and a scalar, k, which is a small number multiplied by the exchange quantity, a decision variable of agents' problems. The first constant value, r_{ij}, can be interpreted as a per unit base price, and the resource price is adjusted based on the exchange quantity and the second constant value, the scalar k. When k is positive/negative, the unit price of resources increases/decreases in exchange quantity.

$$f_{ij}(t, \Phi) = f_{ij}(t, r, k) = r_{ij} + k \times t_{ij} \tag{2.5}$$

Price discrimination: Note that in the price function (2.5), the base price of the same resource can be different for different agents. In other words, agents may pay/receive different amounts for acquiring/selling the same quantity of a resource. This is known as price discrimination. We study a similar price function, presented in (2.6), without price discrimination by removing the j–index from the constant part of the augmented price function, and we show that it can be as effective as the original function, given certain conditions.

$$f'_{ij}(t, \Phi) = f'_{ij}(t, r, k) = r_i + k \times t_{ij} \tag{2.6}$$

Agents' objective function: Agents can make profit through two channels: production and exchange. The profit through producing and selling their products is simply $\pi_j x_j$. The net profit through exchanging resources is $-\sum_i (r_{ij} + k \times t_{ij})\, t_{ij}$. The objective function of agent j, denoted by z_j, is the following:

$$z_j = \pi_j x_j - \sum_i (r_{ij} + k \times t_{ij})\, t_{ij} \tag{2.7}$$

Total supply and demand of each resource: After collecting the exchange values from agents' problems, we can calculate the total amount of resource i requested by agents to buy (total demand, denoted by d_i), as well as the total amount of resource i planned for selling (total supply, denoted by s_i). These two parameters are defined in (2.8) and (2.9).

$$d_i = \sum_j \max\{t_{ij}, 0\} \tag{2.8}$$

$$s_i = \sum_j \max\{-t_{ij}, 0\} \tag{2.9}$$

Buying percentage / selling percentage: If the total demand and supply of a resource do not match, not all buying/selling requests can be accommodated within the alliance, and there needs to be at least one agent who uses spot market option. We assume that in case of any shortage of supply or demand, all buyer/seller agents will be able to execute the same proportion of their buying/selling requests within the alliance, and they all have to resort to the spot market for the remaining amounts. To this end, we define the following two parameters:

- selling percentage (ρ_i^s): the percentage of any resource i selling request that can be sold within the alliance (i.e., to buyer agents);

- buying percentage (ρ_i^b): the percentage of any resource i buying request that can be bought from the agents in the alliance.

The mathematical definitions of these two percentage parameters are given in (2.10):

$$\rho_i^s = \begin{cases} 1, & \text{if } d_i \geq s_i \\ \frac{d_i}{s_i}, & \text{otherwise} \end{cases} \qquad \rho_i^b = \begin{cases} 1, & \text{if } d_i \leq s_i \\ \frac{s_i}{d_i}, & \text{otherwise} \end{cases} \qquad \forall i \tag{2.10}$$

Spot market prices: In case of any mismatch between the total supply and demand of any of the resources, agents will use the spot market to satisfy any buying/selling requests that remain unmatched by other agents. We assume simple linear prices for the spot market options. The two spot market price parameters are the following:

- p_i^b = buying price of resource i, (an agent pays $p_i^b \times Q$ to buy Q units of resource i from the spot market);

- p_i^s = selling price of resource i, (an agent receives $p_i^s \times Q$ by selling Q units of resource i in the spot market).

In order for the alliance of agents to be stable, the resource price function announced by the central planner needs to be better than the spot market prices, to avoid a scenario in which agents prefer the spot market over trading with other agents. We model this as a constraint for the pricing mechanism design problem.

Spot market decision variables: While solving their profit maximizing problems, agents can potentially combine inside–alliance exchanges with spot market trades. Depending on spot market prices, this option may or may not be optimal. In order to capture any trades with the spot market, we introduce the following two decision variables:

- λ_{ij} = amount of resource i that agent j buys from the spot market at price p_i^b;
- μ_{ij} = amount of resource i that agent j sells to the spot market at price p_i^s.

Settlement process: In case of any mismatch between total supply and total demand of any resource, the central planner updates the selling and buying quantities of each agent, determining what proportion of agents' buying/selling requests can be satisfied within the alliance, and for how much they need to rely on the spot market. For instance, if the total supply and the total demand for the resource $\hat{i}$ are 15 and 20, respectively, then the buying percentage is $\rho_{\hat{i}}^b = \frac{15}{20} = 0.75$. Now, if an agent has placed a buying order of 4 units of $\hat{i}$, she will only be able to buy 3 out of 4 units from other agents, and for the remaining 1 unit, she has to buy from the spot market with a price tag of $p_{\hat{i}}^b$. Therefore, her total payment for this 4–unit purchase would be $\Big((r_{\hat{i}j} + k \times 3) * 3\Big) + \Big(p_{\hat{i}}^b \times 1\Big)$.

Post–settlement exchange parameters: We define a new set of parameters to capture the amount of original exchange requests assigned to both the alliance and the spot market. We define non–negative variables, as opposed to the unrestricted in sign exchange variables (t_{ij}) variables.

- θ_{ij}^{in} = the amount of resource i that agent j is supposed to buy from other agents (i.e., from **in**side the alliance);
- γ_{ij}^{in} = the amount of resource i that agent j is supposed to sell to other agents (i.e., to agents **in**side the alliance);
- θ_{ij}^{out} = the amount of resource i that agent j is supposed to buy from the spot market (i.e., from **out**side the alliance);

- γ_{ij}^{out} = the amount of resource i that agent j is supposed to sell to the spot market (i.e., to **out**side the alliance);

<u>**Partitioning the set of agents**</u>: Given values of the inside the alliance exchange variables (i.e., $\theta_{ij}^{\text{in}}, \gamma_{ij}^{\text{in}}$), we partition the set of agents for each resource i into three categories:

- $J_i^b = \{j : \theta_{ij}^{\text{in}} > 0\}$, the set of agents who buy resource i from other agents;
- $J_i^s = \{j : \gamma_{ij}^{\text{in}} > 0\}$, the set of agents who sell resource i to other agents;
- $J_i^0 = \{j : \theta_{ij}^{\text{in}} = \gamma_{ij}^{\text{in}} = 0\}$, the set of agents who neither buy nor sell resource i, and do not exchange resource i with other agents.

<u>**Agents' profit after the settlement process**</u>: Agent j's final profit which is computed after the settlement process and possible trades with the spot market is denoted by w_j and is calculated as follows:

$$\begin{aligned} w_j = \pi_j x_j - \sum_i \left(\left(r_{ij} + k\theta_{ij}^{\text{in}} \right) \theta_{ij}^{\text{in}} + \left(p_i^b \times (\theta_{ij}^{\text{out}} + \lambda_{ij}) \right) \right) \\ + \sum_i \left(\left(r_{ij} - k\gamma_{ij}^{\text{in}} \right) \gamma_{ij}^{\text{in}} + \left(p_i^s \times (\gamma_{ij}^{\text{out}} + \mu_{ij}) \right) \right) \end{aligned} \tag{2.11}$$

<u>**Central planner's profit**</u>: Because of the nonlinear price function, there can be disparity between the transaction value of sellers and buyers. In other words, according to the quadratic price function (2.5), for any exchange between a pair of agents, the amount that the seller is supposed to receive is not necessarily same as the amount the buyer pays. This difference can be both negative and positive, meaning the central planner is either making profit or paying subsidy to the alliance. The net profit amount of the central planner is denoted by w_0, and its formulation is given by :

$$w_0 = \sum_{i,j} \left(\left((r_{ij} - k \times \gamma_{ij}^{\text{in}})\, \gamma_{ij}^{\text{in}} \right) - \left((r_{ij} + k \times \theta_{ij}^{\text{in}})\, \theta_{ij}^{\text{in}} \right) \right) \tag{2.12}$$

Aggregate profit in the decentralized problem: Once we add agents' profits as well as the central planner's profit, we can compute the aggregate profit in the decentralized problem, denoted by w. Since this aggregate profit function, as well as agents' and central planner's profit functions, depend on our choice of the price function and the agents' production and exchange quantities, we explicitly show this dependency by adding in the (Φ, x, t) parameters as arguments of these functions. Equation (2.13) shows the way the aggregate profit is calculated using agents' and the central planner's profit values.

$$w(\Phi, x, t) = \sum_j w_j(\Phi, x, t) + w_0(\Phi, x, t) \tag{2.13}$$

Efficiency ratio: Now that we have calculated final profit value in both centralized problem (i.e., z_c) and in the decentralized problem (i.e., $w(\Phi, x, t)$), we can take the ratio between the decentralized profit and the centralized profit to find the *efficiency ratio*, denoted by $\delta(\Phi)$. We can calculate a distinct efficiency ratio for each choice of the price function, that is why we explicitly show this dependency on price parameters, Φ. The minimization in the numerator refers to the worst–case analysis of the efficiency ratio, and we consider those (x, t) solutions which are optimal for each agent j, while the combination of them can potentially lead to a less than optimal (i.e., the profit in the first best solution) profit value. Let set $\mathbf{H}$ denote the set of optimal solutions (x, t) for all individual agents' problems.

$$\delta(\Phi) = \frac{\min_{(x,t)\in\mathbf{H}} \; w^*(\Phi, x, t)}{z_c^*} \tag{2.14}$$

The main research question in this chapter is to design price parameters, Φ, in a way that the worst case of the aggregate profit (i.e., $w(\Phi, x, t)$) is as close to the central planner's profit (i.e., z_c) as possible. That is, the goal is to find a resource price function to maximize the efficiency ratio, $\delta(\Phi)$. Mathematically speaking, the main problem in this study is to solve the following maximization problem:

$$\delta^* = \max_{\Phi} \ \delta(\Phi) \tag{2.15}$$

Our goal is to analyze the δ^* in depth, and possibly prove that it is exactly equal to 1, subject to certain conditions. The key research question here is whether δ^* is guaranteed to be one, without price discrimination, and if so, what are the characteristics of the optimal resource price function.

2.3.3 Analysis

In this section, we introduce the mathematical models for the central planner's problem, agents' problems, and the resource pricing problem. We then present theoretical results on the conditions under which the efficiency ratio of one can be achieved without price discrimination.

2.3.3.1 Central Planner's Problem

The model M1, (2.16)–(2.18), formulates the central planner's problem with production variables only. The resource constraint (2.17) ensures that the total amount of each resource i used by all the agents cannot exceed the total endowment of resource i within the alliance. The non–negative dual variables p_i are also shown in the formulation.

$$[\text{M1}]\ z_c = \max \quad \sum_j \pi_j x_j^c \tag{2.16}$$

$$\text{s.t.} \quad \sum_j a_{ij} x_j^c \leq b_i, \quad (p_i \geq 0) \qquad \forall i \in R \tag{2.17}$$

$$x_j^c \geq 0, \qquad \forall j \in J \tag{2.18}$$

We develop an equivalent formulation of the central planner's problem which can be useful in our analysis later.

Lemma 2.1. *The model M2 shares the same feasible region and the same set of optimal solutions as the model M1.*

$$\text{[M2] } z_c = \max \quad \sum_j \pi_j x_j^c \tag{2.19}$$

$$s.t. \quad a_{ij} x_j^c \leq b_{ij} + t_{ij}^c, \quad (q_{ij} \geq 0) \qquad \forall i \in R, \forall j \in J \tag{2.20}$$

$$\sum_j t_{ij}^c = 0, \quad (h_i) \qquad \forall i \in R \tag{2.21}$$

$$x_j^c \geq 0, \qquad \forall j \in J \tag{2.22}$$

Proof. First, we show that given any feasible solution $(\bar{x}_j^c, \bar{t}_{ij}^c)$ in $M2$, we can always construct a feasible solution in $M1$. All we need to do is to sum over the set of agents in constraint (2.20) of the model $M2$, and since the summation of all the exchange variables is zero due to (2.21), the remaining inequality is exactly equal to the resource constraint (2.17) of the model $M1$. With the same resource constraint and the same non–negativity constraint, we conclude that the feasible solution $(\bar{x}_j^c, \bar{t}_{ij}^c)$ always maps to a feasible solution in $M1$. Now, we need to show that given any feasible solution $(\bar{x}_j^c)$ in $M1$, we can construct a feasible solution in $M2$. To this end, we devise Algorithm 1, which takes the production quantities of the agents, and compares the required resources for each agent with the corresponding endowment amount of each agents. The algorithm then assigns exchange values to the agents in a minimal way, while satisfying the condition (2.21).

We have proved that the two optimization problems share the same feasible region. Now with the same feasible region and the same objective function (i.e., $max \sum_j \pi_j x_j^c$), the set of optimal solutions will also be the same, as they both are basically modeling the same problem. $\square$

Algorithm 1 Mapping the solution ($\bar{x}^c$) in M1 to a solution (x^c, t^c) in M2

$x^c \leftarrow \bar{x}^c$
for $i \in R$ **do**
  $s_i \leftarrow 0$
  $J_i^s \leftarrow \emptyset$
  for $j \in J$ **do**
    $t_{ij}^c \leftarrow \max\{0, a_{ij}x_j^c - b_{ij}\}$
    $s_i \leftarrow s_i + t_{ij}^c$
    if $a_{ij}x_j^c < b_{ij}$ **then**
      $J_i^s \leftarrow J_i^s \cup j$
    end if
  end for
end for
for $i \in R$ **do**
  for $j \in J_i^s$ **do**
    $t_{ij}^c \leftarrow \min\{s_i, b_{ij} - a_{ij}x_j^c\}$
    $s_i \leftarrow s_i - t_{ij}^c$
    if $s_i = 0$ **then**
      break
    end if
  end for
end for

Proposition 2.1. *The following relationship holds for the optimal dual variables in $M1$ and $M2$:*

$$p_i^* = h_i^* = q_{ij}^* \qquad \forall i \in R, \forall j \in J \tag{2.23}$$

Proof. First, let us take the dual of $M2$ which will help us unpack some of the equations. Let us denote this dual problem by $D2$.

$$[\text{D2}]\ z_c = \min \quad \sum_{i,j} b_{ij}q_{ij} \tag{2.24}$$

$$\text{s.t.} \quad \sum_i a_{ij}q_{ij} \geq \pi_j, \quad (x_j \geq 0) \qquad \forall j \in J \tag{2.25}$$

$$-q_{ij} + h_i = 0, \quad (t_{ij}) \qquad \forall i \in R, \forall j \in J \tag{2.26}$$

$$q_{ij} \geq 0, \qquad \forall i \in R, \forall j \in J \tag{2.27}$$

Constraint (2.26) of the dual problem $D2$ gives us the equality of q_{ij} and h_i variables in (2.23). Now, we replace q_{ij} with its equivalent value, h_i, in the dual problem $D2$. The result, denoted by $D1$, is the following model (where here we have used the fact that $b_i = \sum_j b_{ij}$):

$$
\begin{aligned}
[\text{D1}]\ z_c = \min \quad & \sum_i b_i h_i \\
\text{s.t.} \quad & \sum_i a_{ij} h_i \geq \pi_j, && \forall j \in J \\
& h_i \geq 0, && \forall i \in R
\end{aligned}
$$

Note that problem $D1$ is precisely the dual problem of the $M1$ formulation, thus share the same feasible and optimal solutions (i.e., $p_i^* = h_i^*, \forall i \in R$), and that proves the equality of p_i and h_i variables for all resources. It is not difficult to observe that given the two proven equations, $p_i = q_{ij}$ also holds, thus completing the proof for (2.23). □

Now that we have observed the equivalence of the two optimization models, $M1$ and $M2$, and we have established the relationship between their optimal dual variables, we know that the optimal profit is the same in $M1$ and $M2$, and that it is equal to z_c. This profit value is the denominator of the key metric in this study, efficiency ratio δ. The rest of the chapter will be spent on analyzing the agents' problems.

2.3.3.2 Agents' Problems

With the definitions of the augmented resource price function, agents' objective function and resource constraints discussed in section 2.3.2, we now introduce the full form of agents' profit maximizing problems, denoted by P_j.

$$
[P_j]\ z_j = \max \quad \pi_j x_j - \sum_i \Big((r_{ij} + k \times t_{ij}) t_{ij} \Big) \tag{2.28}
$$

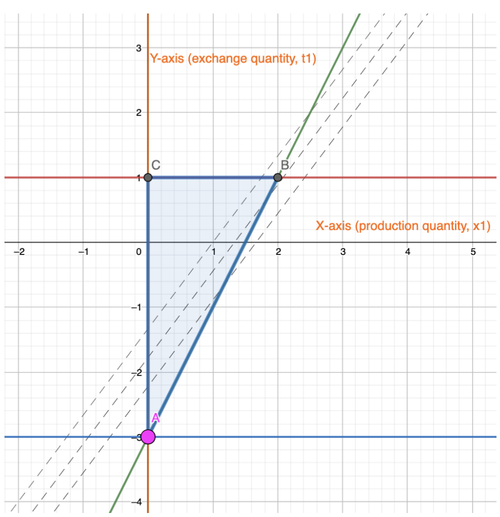
$$\begin{aligned}
\text{s.t.} \quad & a_{ij}x_j \leq b_{ij} + t_{ij}, \quad (y_{ij} \geq 0) && \forall i \in R, \forall j \in J && (2.29)\\
& t_{ij} \leq \bar{b}_{ij}, \quad (v_{ij} \geq 0) && \forall i \in R, \forall j \in J && (2.30)\\
& -t_{ij} \leq b_{ij}, \quad (u_{ij} \geq 0) && \forall i \in R, \forall j \in J && (2.31)\\
& x_j \geq 0, && \forall j \in J && (2.32)
\end{aligned}$$

The objective function (2.28) has two components: profit through production and net cost through exchange of resources with other agents. Constraint (2.29) ensures that the resource usage is bounded by the endowment amount, adjusted by exchange quantities. The two bounds on the exchange variable, (2.30) and (2.31), make sure that agents don't buy more than what is available in the market, and they don't sell more than their own endowment values. The same optimization problem, adding the spot market option to the problem is the following:

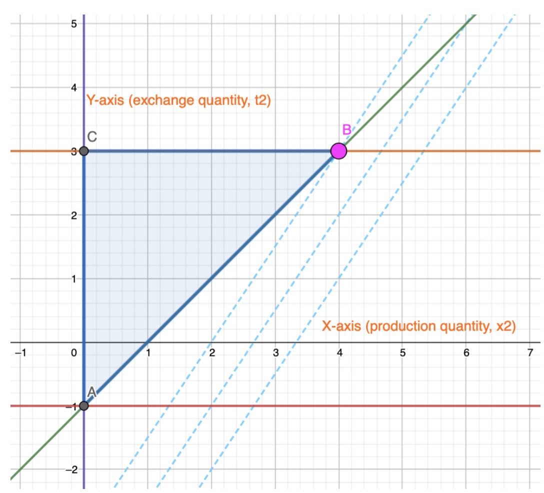
$$\begin{aligned}
[P_j^{\lambda\mu}]\ z_j^{\lambda\mu} = \max \quad & \pi_j x_j - \sum_i \left(\Big((r_{ij} + k \times t_{ij}) t_{ij} \Big) + \Big(p_i^b \lambda_{ij} - p_i^s \mu_{ij} \Big) \right) \\
\text{s.t.} \quad & a_{ij}x_j \leq b_{ij} + t_{ij} + \lambda_{ij} - \mu_{ij}, \quad (y_{ij} \geq 0) && \forall i \in R, \forall j \in J\\
& t_{ij} \leq \bar{b}_{ij}, \quad (y_{ij} \geq 0) && \forall i \in R, \forall j \in J\\
& -t_{ij} \leq b_{ij}, \quad (y_{ij} \geq 0) && \forall i \in R, \forall j \in J\\
& x_j, \lambda_{ij}, \mu_{ij} \geq 0, && \forall i \in R, \forall j \in J
\end{aligned}$$

We discussed in section 2.3.2 why the augmented price function is an appropriate choice for the coordination problem under investigation in this chapter. One of the reasons why it is a great choice is the fact that it covers a wide range of pricing schemes studied in the literature, including linear dual pricing. When the constant k is zero, the price function f reduces to a linear price function. Even though linear dual prices have been studied extensively in the literature, it is well–known that suboptimization by agents on the basis of linear dual prices does not lead to globally optimal solutions, in general, due to lack of coordination. The main reason for the failure of the linear pricing strategy is the possibility of multiple optimal solutions in agents' problems which leads to the lack of coordination. We illustrate this point using a small 2–agent 1–resource numerical example. Imagine there are two agents, using a single resource to produce their products. Agent 1 makes \$6 by selling one unit of her product, while agent 2 makes \$3. Agents 1 and 2 have been endowed with 3 and 1 units of resource, respectively. The resource usage rate for the two agents are 2 units per agent 1's product and 1 unit per agent 2's product. The two individual optimization problems are shown below, followed by their feasible region in Figure 2.3.

$$
\begin{aligned}
[P_1]\ z_1 = \max \quad & 6x_1 - pt_1 \\
\text{s.t.} \quad & 2x_1 \leq 3 + t_1 \\
& -3 \leq t_1 \leq 1 \\
& x_1 \geq 0
\end{aligned}
\qquad\qquad
\begin{aligned}
[P_2]\ z_2 = \max \quad & 3x_2 - pt_2 \\
\text{s.t.} \quad & x_2 \leq 1 + t_2 \\
& -1 \leq t_2 \leq 3 \\
& x_2 \geq 0
\end{aligned}
$$



(a) Agent 1



(b) Agent 2

Figure 2.3: Feasible regions for two agents' optimization problems

According to Figure 2.3, both agents' problems have three corner points (A,B,C), and depending on the price p, which changes the direction and the level curves of the objective function, the optimal solution can be either of these three points. Assuming the price is reasonable and the profit does not fall below zero, the optimal solution will be one of the two corner points (A,B) for both agents. It is not difficult to see that if $p \geq 3$, point $A : (x_1 = 0, t_1 = -3)$ is the optimal corner point for agent 1, and if $p \leq 3$, point $B : (x_1 = 2, t_1 = 1)$ is the optimal solution. Note that both corner points are optimal when $p = 3$ which is an indication of multiple optimal solutions. Similar argument can be made for agent 2. Point $A : (x_2 = 0, t_2 = -1)$ is optimal for agent 2 when $p \geq 3$, and point $B : (x_2 = 4, t_2 = 3)$ is optimal when $p \leq 3$. Both points A and B are optimal when $p = 3$. When considering both problems at the same time, agents' solutions do not match when $p \neq 3$, and when $p = 3$, the worst possible case still is "no coordination" which results in an efficiency ratio less than one. This small example illustrates why linear pricing without price discrimination may not be able to coordinate the agents. Note that if we are allowed to choose different price values for agents 1 and 2, we can certainly direct both problems into a coordinating solution.

Nonlinear pricing however can ensure uniqueness in agents' problems. Having a unique solution in each agent's problem, the main concern remains to be the value of those solutions and whether they are in agreement with the centralized solution. Moreover, even if the solutions do match that of the central planner's problem, whether there is price discrimination or the need for subsidization by the central planner. Convex optimization problems with strictly concave objective function and a convex feasible region yield a unique optimal solution. Agents' problems are convex quadratic programming problems, as the objective function is strictly convex quadratic (when $k > 0$), constraints are linear, and the decision variables are continuous. As a result, agents' problems have unique optimal solutions with a nonlinear price function.

2.3.3.3 Modeling the Resource Pricing Problem

Now that we have a tool to produce unique optimal solutions in agents' problems, we need to revisit the main research question of this chapter, namely designing the resource price function with the goal of achieving the maximum efficiency ratio. Mathematically speaking, the main problem is to find a price function $\Phi = (r, k)$ such that the efficiency ratio $\delta(\Phi)$ is maximized.

$$\delta^* = \max_{\Phi} \ \delta(\Phi) = \max_{\Phi} \ \frac{\min_{(x,t)\in \mathbf{H}} \ w(\Phi, x, t)}{z_c} \tag{2.33}$$

According to the expanded formula of the optimal efficiency ratio, for any choice of the price function Φ we can compute an efficiency ratio $\delta(\Phi)$. The goal is to find a price function that maximizes the efficiency ratio. The denominator of the efficiency ratio does not depend on our choice of price function. The numerator however depends on both the price function, as well as the production and exchange values. The minimization in the numerator is important only when there are multiple optimal solutions in some of agents' problem. Finding the worst overall solution of the decentralized problem is a very challenging problem with both agents' decisions (i.e., production and exchange quantities) and the linear resource prices as decision variables of the problem. In order to write down the full minimization problem in the numerator of (2.33), we need to write the optimality conditions for all agents. We use *Karush–Kuhn–Tucker (KKT) conditions* for this purpose. The union of all KKT conditions for all agents provides a feasibility problem. We then add the total aggregate profit of the decentralized problem, $w(\Phi, x, t)$, to the objective function, and minimizing this quantity results in the worst overall solution in the decentralized problem. We should note that calculating $w(\Phi, x, t)$ is a two–stage process which involves solving the feasibility problem in the first stage, and then executing the settlement process in the second stage.

In the following subsections, first, we derive our results for linear prices ($k = 0$), and then we extend the results to the general nonlinear prices ($k > 0$).

2.3.3.4 Worst Decentralized Solution with Linear Resource Prices

Using linear prices r_{ij} (only the constant part, without the variable part), agents' problems are the following:

$$
\begin{aligned}
[P_j]\ z_j = \max \quad & \pi_j x_j - \sum_i r_{ij} t_{ij} && && (2.34)\\
\text{s.t.} \quad & a_{ij} x_j \leq b_{ij} + t_{ij}, \quad (y_{ij} \geq 0) && \forall i \in R, \forall j \in J && (2.35)\\
& t_{ij} \leq \bar{b}_{ij}, \quad (v_{ij} \geq 0) && \forall i \in R, \forall j \in J && (2.36)\\
& -t_{ij} \leq b_{ij}, \quad (u_{ij} \geq 0) && \forall i \in R, \forall j \in J && (2.37)\\
& x_j \geq 0, && \forall j \in J && (2.38)
\end{aligned}
$$

We write the KKT conditions for each agent $j \in J$, and then take the union of all the individual KKT conditions, and form the aforementioned feasibility problem.

1. **Stationarity** : ($-\nabla$ max. objective $+ \sum$ dual variables $\times \nabla$ constraints $= 0$)

$$x_j \longrightarrow -\pi_j + \sum_i y_{ij} a_{ij} = 0, \tag{2.39}$$

$$t_{ij} \longrightarrow r_{ij} - y_{ij} - u_{ij} + v_{ij} = 0, \quad \forall i \in R$$

2. **Primal Feasibility** :

$$
\begin{aligned}
& a_{ij} x_j \leq b_{ij} + t_{ij}, \quad \forall i \in R\\
& t_{ij} \leq \bar{b}_{ij}, \quad \forall i \in R\\
& -t_{ij} \leq b_{ij}, \quad \forall i \in R\\
& x_j \geq 0
\end{aligned}
\tag{2.40}
$$

3. Dual Feasibility :

$$y_{ij}, u_{ij}, v_{ij} \geq 0, \quad \forall i \tag{2.41}$$

4. Complementary Slackness :

$$\begin{aligned} y_{ij} \times \Big(b_{ij} + t_{ij} - a_{ij}x_j\Big) \ &= \ 0, \quad \forall i \\ u_{ij} \times \Big(b_{ij} + t_{ij}\Big) \ &= \ 0, \quad \forall i \\ v_{ij} \times \Big(\bar{b}_{ij} - t_{ij}\Big) \ &= \ 0, \quad \forall i \end{aligned} \tag{2.42}$$

Taking the union of KKT conditions for each agent j and using (2.13) in the objective function, we present the two–stage model in a single formulation, denoted by (P–LIN). In order to better observe the complexity of the problem, decision variables are shown with a different color. Variables $(w_j, w_0, r_{ij}, y_{ij}, u_{ij}, v_{ij}, x_j, t_{ij}, d_i, s_i, \rho_i^s, \rho_i^b, \theta_{ij}^{\text{in}}, \theta_{ij}^{\text{out}}, \gamma_{ij}^{\text{in}}, \gamma_{ij}^{\text{out}})$ are all the decision variables in problem [P–LIN].

$$\begin{aligned} \text{[P–LIN]} \ \min \quad & \sum_j w_j + w_0 \\ \text{s.t.} \quad & -\pi_j + \sum_i a_{ij} y_{ij} = 0, && \forall j \in J \\ & r_{ij} - y_{ij} - u_{ij} + v_{ij} = 0, && \forall i \in R, \forall j \in J \\ & a_{ij} x_j \leq b_{ij} + t_{ij}, && \forall i \in R, \forall j \in J \\ & t_{ij} \leq \bar{b}_{ij}, && \forall i \in R, \forall j \in J \\ & -t_{ij} \leq b_{ij}, && \forall i \in R, \forall j \in J \\ & x_j \geq 0, && \forall j \in J \\ & y_{ij}, u_{ij}, v_{ij} \geq 0, && \forall i \in R, \forall j \in J \\ & y_{ij} \times \Big(b_{ij} + t_{ij} - a_{ij} x_j\Big) \ = \ 0, && \forall i \in R, \forall j \in J \end{aligned}$$

$$u_{ij} \times \left(b_{ij} + t_{ij}\right) = 0, \qquad \forall i \in R, \forall j \in J$$

$$v_{ij} \times \left(\bar{b}_{ij} - t_{ij}\right) = 0, \qquad \forall i \in R, \forall j \in J$$

$$d_i = \sum_j \max\{t_{ij}, 0\}, \qquad \forall i \in R$$

$$s_i = \sum_j \max\{-t_{ij}, 0\} \qquad \forall i \in R$$

$$\rho_i^s = \begin{cases} 1, & \text{if } d_i \geq s_i \\ \frac{d_i}{s_i}, & \text{otherwise} \end{cases} \qquad \rho_i^b = \begin{cases} 1, & \text{if } d_i \leq s_i \\ \frac{s_i}{d_i}, & \text{otherwise} \end{cases} \qquad \forall i \in R$$

$$\theta_{ij}^{\text{in}} = \rho_i^b \times \max\{0, t_{ij}\} \qquad \forall i \in R, \forall j \in J$$

$$\theta_{ij}^{\text{out}} = (1 - \rho_i^b) \times \max\{0, t_{ij}\} \qquad \forall i \in R, \forall j \in J$$

$$\gamma_{ij}^{\text{in}} = \rho_i^s \times \max\{0, -t_{ij}\} \qquad \forall i \in R, \forall j \in J$$

$$\gamma_{ij}^{\text{out}} = (1 - \rho_i^s) \times \max\{0, -t_{ij}\} \qquad \forall i \in R, \forall j \in J$$

$$w_j = \pi_j x_j - \sum_i \left(r_{ij} \times \theta_{ij}^{\text{in}} + p_i^b \times \theta_{ij}^{\text{out}}\right)$$

$$+ \sum_i \left(r_{ij} \times \gamma_{ij}^{\text{in}} + p_i^s \times \gamma_{ij}^{\text{out}}\right) \qquad \forall j \in J$$

$$w_0 = \sum_{i,j} r_{ij}(\gamma_{ij}^{\text{in}} - \theta_{ij}^{\text{in}})$$

The problem formulated in (P–LIN) is difficult to solve, as both the constraints and the objective function are nonlinear. The objective function is the aggregate profit post–settlement. The first ten constraints are the KKT conditions in the linear pricing case written for all agents. The next four constraints define the total supply (s_i) and the total demand (d_i) of each resource, as well as selling (ρ_i^s and buying (ρ_i^b) percentages, all as functions of the exchange variable (t_{ij}) found in the KKT conditions. Both exchange variables and the two percentage variables are then used to determine the proportion of the exchange requests happening within the alliance, and likewise the proportion of such exchanges outsourced to the spot market. These quantities are being captured by the following four decision variables:

$\theta_{ij}^{\text{in}}, \theta_{ij}^{\text{out}}, \gamma_{ij}^{\text{in}}, \gamma_{ij}^{\text{out}}$. Finally the true profit of each agent, w_j, as well as the central planner's profit, w_0, are calculated using updated exchange values which then is used in the objective function. Note that the central planner's profit (or subsidy) is zero only when there is no price discrimination. Linear resource prices with price discrimination (i.e., using a different resource price r_{ij} for each agent) can always achieve the efficiency ratio of one. Based on extensive numerical analysis that is omitted here for brevity (details, available in section 2.4), it is highly likely that linear resource prices with price discrimination (i.e., using a different resource price r_{ij} for each agent) can always achieve the efficiency ratio of one.

In essence, there is a tradeoff between linear prices with price discrimination and linear prices without price discrimination. The main advantage of the former is achieving the best efficiency ratio (i.e., $\delta^* = 1$), while it increases the central planner's potential subsidy amount. The main advantage of the latter case is eliminating the subsidy, while having a suboptimal solution in terms of the efficiency ratio (i.e., $\delta^* \leq 1$).

2.3.3.5 Resource Pricing Problem with Nonlinear Prices

As discusses previously in this section, assuming a unique solution in all agents' problems, which can be attained using the augmented price function (2.5) when $k > 0$, the formula for the efficiency ratio reduces to the following, where agents' problems have unique optimal solutions; thus the minimization over multiple choice of the optimal solutions (x^*, t^*) in $w(\Phi, x, t)$ is omitted:

$$\delta^* = \max_{\Phi} \ \delta(\Phi) = \max_{\Phi} \ \frac{w(\Phi, x, t)}{z_c} \tag{2.43}$$

Jennergren (1972) has shown that the augmented price function of the form (2.5) achieves the efficiency ratio of one, by ensuring that the central planner's optimal solution is also optimal for all the agents. Jennergren also characterizes the optimal price parameters (i.e., r_{ij}, k)

which are based on the optimal primal (t_{ij}^{c*}) and dual (p_i^*) values of the central planner's problem (M2). The optimal resource price function according to Jennergren (1972, 1973) is the following equation:

$$\begin{aligned} f_{ij}(r,k,t) &= r_{ij} + k \times t_{ij}, \\ r_{ij} &= p_i^* - 2k \times t_{ij}^{c*} \end{aligned} \tag{2.44}$$

For a ($k > 0$), but sufficiently small, price function (2.44) achieves an efficiency ratio of one (Jennergren 1972). There are mainly two drawbacks with Jennergren's approach. The first drawback is with price discrimination. According to Jennergren's approach, agents can potentially have different price functions for the same resource. The second problem with Jennergren's approach is the fact that the central planner might have to subsidize some exchanges. Therefore, achieving full coordination costs the central planner. First, we propose a method that achieves the efficiency ratio of one for the single–resource case with an extension to the multiple–resource case as a conjecture, without price discrimination, using the resource price function introduced in (2.45), below.

$$f_{ij}(r,k,t) = r_i + k \times t_{ij} \tag{2.45}$$

We then analyze the subsidy amount for newly proposed resource price function. First, let us revisit the agents' profit maximizing problems with nonlinear price function of the form (2.45):

$$[P_j]\ z_j = \max \quad \pi_j x_j - \sum_i \Big(r_i + k \times t_{ij}\Big) t_{ij} \tag{2.46}$$

$$\text{s.t.} \quad a_{ij} x_j \le b_{ij} + t_{ij}, \quad (y_{ij} \ge 0) \qquad \forall i \in R, \forall j \in J \tag{2.47}$$

$$t_{ij} \le \bar{b}_{ij}, \quad (v_{ij} \ge 0) \qquad \forall i \in R, \forall j \in J \tag{2.48}$$

$$-t_{ij} \le b_{ij}, \quad (u_{ij} \ge 0) \qquad \forall i \in R, \forall j \in J \tag{2.49}$$

$$x_j \geq 0, \qquad \forall j \in J \tag{2.50}$$

For simplicity, we start the discussion with an instance of the problem having only one resource (and multiple agents).

Theorem 2.1. *With the resource price function* (2.45)*, the efficiency ratio is always one in single–resource instances of the problem.*

Proof. With only a single resource, we can simplify agent's problems and formulate them using only the exchange variables. This property holds because all the resource constraints are binding for all the agents, and we can model the production variable as a function of the exchange variables. Model $[P_j^1]$ shows agent j's problem with a single resource.

$$[P_j^1]\ z_j^1 = \max \quad \pi_j x_j - (r + k \times t_j)t_j \tag{2.51}$$

$$\text{s.t.} \quad a_j x_j = b_j + t_j, \quad (y_j \geq 0) \qquad \forall j \in J \tag{2.52}$$

$$t_j \leq \bar{b}_j, \quad (v_j \geq 0) \qquad \forall j \in J \tag{2.53}$$

$$-t_j \leq b_j, \quad (u_j \geq 0) \qquad \forall j \in J \tag{2.54}$$

$$x_j \geq 0, \qquad \forall j \in J \tag{2.55}$$

Writing the same model using only the exchange variables, $[P_j^1]$ for agent j is formulation (2.56)–(2.58). In this transformation, we replaced variable x_j with $\frac{b_j+t_j}{a_j}$, obtained from constraint (2.52).

$$[P_j^1]\ z_j^1 = \max \quad t_j \times (\frac{\pi_j}{a_j} - r) - k \times t_j^2 \tag{2.56}$$

$$\text{s.t.} \quad t_j \leq \bar{b}_j, \quad (v_j \geq 0) \qquad \forall j \in J \tag{2.57}$$

$$-t_j \leq b_j, \quad (u_j \geq 0) \qquad \forall j \in J \tag{2.58}$$

Now, let us write the KKT conditions for agent j's problem.

1. **Stationarity** : $(-\nabla\,\text{max. objective} + \sum \text{dual variables} \times \nabla\,\text{constraints} = 0)$

$$t_j \longrightarrow -\frac{\pi_j}{a_j} + r + 2kt_j + v_j - u_j = 0 \tag{2.59}$$

2. **Primal Feasibility** :

$$\begin{aligned} &t_j \le \bar{b}_j, \\ &-t_j \le b_j. \end{aligned} \tag{2.60}$$

3. **Dual Feasibility** :

$$u_j, v_j \ge 0. \tag{2.61}$$

4. **Complementary Slackness** :

$$\begin{aligned} u_j \times \Big(b_j + t_j\Big) &= 0, \\ v_j \times \Big(\bar{b}_j - t_j\Big) &= 0, \end{aligned} \tag{2.62}$$

Combining all the KKT conditions for all the agents (set J), the feasibility problem (P^1–NL) solves the coordinating problem. To prove Theorem 2.1, it suffices to show that problem (P^1–NL) is always feasible. Feasibility of this problem implies that the resource price function (2.45) always achieves an efficiency ratio of one for single–resource instance. Variables (k, r, u_j, v_j, t_j) are the decision variables, shown in different color in problem [P^1–NL].

$$\begin{aligned} [\text{P}^1\text{–NL}]\ \ \max\quad & 0 \\ \text{s.t.}\quad & -\frac{\pi_j}{a_j} + r + 2\,k\,t_j - u_j + v_j = 0, \qquad \forall j \in J \end{aligned}$$

$$
\begin{aligned}
& t_j \leq \bar{b}_j, && \forall j \in J \\
& -t_j \leq b_j, && \forall j \in J \\
& u_j, v_j \geq 0, && \forall j \in J \\
& u_j \times \left(b_j + t_j\right) = 0, && \forall j \in J \\
& v_j \times \left(\bar{b}_j - t_j\right) = 0, && \forall j \in J
\end{aligned}
$$

The first constraint in (P^1–NL) suggests a formula for the base price of the resource price function, r:

$$r = \frac{\pi_j}{a_j} - 2\,k\,t_j + u_j - v_j \tag{2.63}$$

We divide the set of all agents J into two subsets: the set of producer agents, denoted by J^+, and the set of non–producer agents, denoted by J^0. We will use a target solution (t) to further simplify the problem (P^1–NL) and to prove that a single r in (2.63) can satisfy the feasibility problem. The exchange quantity of the first best solution can potentially solve the coordinating problem without price discrimination. Let t_j^* denote the central planner's solution for agent j's exchange quantity. In essence, agents either produce a product which may require acquiring more resources from other agents, or they sell all their resource endowments. Given this partitioning, the value of r for each of these two subset of agents is the following:

$$r = \begin{cases} \frac{\pi_j}{a_j} - (2\,k\,t_j + v_j), & \text{if } j \in J^+, \\ \frac{\pi_j}{a_j} + (2\,k\,b_j + u_j), & \text{if } j \in J^0. \end{cases} \tag{2.64}$$

Note that the mathematical expression inside the parentheses for both types of agents is non–negative, as t_j is non–negative for producer agents and both dual variables, v_j, u_j, are also non–negative by definition. As a result, we have the freedom to increase or decrease the dual variables without any restriction to make sure the same r value satisfies all the constraints. Therefore, we can conclude that the single–r resource price function always achieves an efficiency of one for single–resource instance. $\square$

According to Theorem 2.1, in a single–resource case, agents' problems can be formulated in a simpler way with fewer decision variables. Moreover, a single base price (i.e., r) unique for all the agents is guaranteed to achieve the efficiency ratio of one. Next, we extend our results to the multiple–resources case.

Based on some structural properties of the KKT conditions (which are explained later in this section) as well as our extensive numerical analysis that is omitted here for brevity (details available from the authors), we observe that it is highly likely that with the resource price function (2.45), the efficiency ratio is one in the multiple–resource multiple–agent instance of the problem, when the solution to the resource allocation problem satisfies the following two conditions:

1. **Minimal exchange**: agents make minimal exchange meaning that the total amount of resources used for production among all agents is exactly equal to the endowment of agents producing plus the amount of resources being exchanged among buyers and sellers.

$$a_{ij}x_j = b_{ij} + t_{ij}, \qquad \forall i \in R,\ j \in J \tag{2.65}$$

2. **Equal division**: for partially–used resources, the total demand from the producer agents is equally divided among non–producer agents. Let $\bar{t}_i$ be that equal amount for resource i. Assume R^p, J^+, J^0 denote the set of partially used resources, producer

agents, and non–producer agents, respectively.

$$\bar{t}_i = \frac{1}{|J^0|} \sum_{j \in J^+} t^*_{ij} \tag{2.66}$$

With a nonlinear price function, agents' problems have unique optimal solutions. The value of the optimal solution for each agent depends on the objective function, more specifically the resource price function. There can be two approaches towards ensuring the best efficiency ratio with a price function that does not differentiate among agents:

Method #1: Similar to the problem (P–LIN) with linear resource price function, we can consider both price parameters, (r_i, k) in this case, as well as agents' decisions (i.e., production and exchange quantities) as decision variables of a problem whose constraint set includes the optimality conditions of all the agents. Since there is no multiple optimal solutions scenario in this case, no objective function is required for the feasibility problem with the optimality conditions, and we can use a simple $(\max 0)$ objective function. Let us review the KKT conditions for each agent $j \in J$, with nonlinear pricing of the form (2.45):

1. **Stationarity** : $(-\nabla \max.\ \text{objective} + \sum \text{dual variables} \times \nabla\, \text{constraints} = 0)$

$$x_j \longrightarrow -\pi_j + \sum_i y_{ij} a_{ij} = 0, \tag{2.67}$$

$$t_{ij} \longrightarrow r_{ij} + 2k t_{ij} - y_{ij} - u_{ij} + v_{ij} = 0, \quad \forall i \in R$$

2. **Primal Feasibility** :

$$\begin{aligned} & a_{ij} x_j \leq b_{ij} + t_{ij}, \quad \forall i \in R \\ & t_{ij} \leq \bar{b}_{ij}, \quad \forall i \in R \\ & -t_{ij} \leq b_{ij}, \quad \forall i \in R \\ & x_j \geq 0 \end{aligned} \tag{2.68}$$

3. Dual Feasibility :

$$y_{ij}, u_{ij}, v_{ij} \geq 0, \quad \forall i \tag{2.69}$$

4. Complementary Slackness :

$$\begin{aligned} y_{ij} \times \Big(b_{ij} + t_{ij} - a_{ij}x_j\Big) &= 0, \quad \forall i \\ u_{ij} \times \Big(b_{ij} + t_{ij}\Big) &= 0, \quad \forall i \\ v_{ij} \times \Big(\bar{b}_{ij} - t_{ij}\Big) &= 0, \quad \forall i \end{aligned} \tag{2.70}$$

Taking the union of KKT conditions for each agent j, we present the two–stage model (i.e., determining the resource price function, and agents' production and exchange decisions) in a single formulation, denoted by (P–NL). In order to better observe the complexity of the problem, similar to (P–LIN), decision variables are shown with a different color. Variables $(w_j, w_0, k, r_{ij}, y_{ij}, u_{ij}, v_{ij}, x_j, t_{ij}, d_i, s_i, \rho_i^s, \rho_i^b, \theta_{ij}^{\text{in}}, \theta_{ij}^{\text{out}}, \gamma_{ij}^{\text{in}}, \gamma_{ij}^{\text{out}})$ are all the decision variables in problem [P–NL].

$$\begin{aligned} \text{[P–NL]} \ \max \quad & 0 & \\ \text{s.t.} \quad & -\pi_j + \sum_i a_{ij} y_{ij} = 0, & \forall j \in J \\ & r_i + 2k t_{ij} - y_{ij} - u_{ij} + v_{ij} = 0, & \forall i \in R, \forall j \in J \\ & a_{ij} x_j \leq b_{ij} + t_{ij}, & \forall i \in R, \forall j \in J \\ & t_{ij} \leq \bar{b}_{ij}, & \forall i \in R, \forall j \in J \\ & -t_{ij} \leq b_{ij}, & \forall i \in R, \forall j \in J \\ & x_j \geq 0, & \forall j \in J \\ & y_{ij}, u_{ij}, v_{ij} \geq 0, & \forall i \in R, \forall j \in J \end{aligned}$$

$$
\begin{aligned}
& y_{ij} \times \Big(b_{ij} + t_{ij} - a_{ij} x_j\Big) \; = \; 0, && \forall i \in R, \forall j \in J \\
& u_{ij} \times \Big(b_{ij} + t_{ij}\Big) \; = \; 0, && \forall i \in R, \forall j \in J \\
& v_{ij} \times \Big(\bar{b}_{ij} - t_{ij}\Big) \; = \; 0, && \forall i \in R, \forall j \in J \\
& d_i = \sum_j \max\{t_{ij}, 0\}, && \forall i \in R \\
& s_i = \sum_j \max\{-t_{ij}, 0\} && \forall i \in R \\
& \rho_i^s = \begin{cases} 1, & \text{if } d_i \geq s_i \\ \frac{d_i}{s_i}, & \text{otherwise} \end{cases} \qquad \rho_i^b = \begin{cases} 1, & \text{if } d_i \leq s_i \\ \frac{s_i}{d_i}, & \text{otherwise} \end{cases} && \forall i \in R \\
& \theta_{ij}^{\text{in}} = \rho_i^b \times \max\{0, t_{ij}\} && \forall i \in R, \forall j \in J \\
& \theta_{ij}^{\text{out}} = (1 - \rho_i^b) \times \max\{0, t_{ij}\} && \forall i \in R, \forall j \in J \\
& \gamma_{ij}^{\text{in}} = \rho_i^s \times \max\{0, -t_{ij}\} && \forall i \in R, \forall j \in J \\
& \gamma_{ij}^{\text{out}} = (1 - \rho_i^s) \times \max\{0, -t_{ij}\} && \forall i \in R, \forall j \in J \\
& w_j = \pi_j x_j - \sum_i \Big((r_i + k\theta_{ij}^{\text{in}})\theta_{ij}^{\text{in}} + p_i^b \times \theta_{ij}^{\text{out}}\Big) \\
& + \sum_i \Big((r_i - k\gamma_{ij}^{\text{in}})\gamma_{ij}^{\text{in}} + p_i^s \times \gamma_{ij}^{\text{out}}\Big) && \forall j \in J \\
& w_0 = \sum_{i,j} \Bigg(\Big((r_i - k\gamma_{ij}^{\text{in}})\,\gamma_{ij}^{\text{in}}\Big) - \Big((r_i + k\theta_{ij}^{\text{in}})\,\theta_{ij}^{\text{in}}\Big)\Bigg)
\end{aligned}
$$

The problem (P–NL), which is an instance of nonlinear programming problems (NLP), is extremely difficult to solve, since the feasible region is non–convex. Therefore, we introduce the second method which solves the pricing problem in a single stage that only involves finding the price function parameters, while production and exchange quantities are pre–determined using the central planner's problem.

Method #2: We take the central planner's optimal solution (x^{c*}, t^{c*}) as our target solution and we use inverse optimization to guarantee the optimality of the central planner's solution for all the agents. We can execute the perturbation scheme of the inverse optimization by changing the parameters of the objective function (e.g., resource price function parameters in agents' problems). According to our extensive numerical study, it is very likely that for a given price function (r_i, k), a primal–dual solution $\Big((x,t) - (y,u,v)\Big)$ satisfies the following three conditions, if and only if (x,t) is an optimal solution to the central planner's problem (i.e., M1 and M2).

1. The primal–dual solution $\Big((x,t) - (y,u,v)\Big)$ satisfies (2.67)–(2.70);
2. For every resource i, total buying and selling orders by the agents match (i.e., market clears);
$$\sum_j t_{ij} = 0$$
3. The aggregate profit is equal to the optimal profit in the central planner's problem (z_c) minus the central planner's net profit (w_o).
$$\sum_j z_j = z_c - w_0$$

Therefore, if there is any solution, optimal for all the agents, while satisfying some of the globally optimal solution conditions (i.e., market clearance and the aggregate profit value), that solution needs to be optimal for the central planner's problem. Therefore, we can assume that the central planner's optimal solution (i.e., x^{c*}, t^{c*}) is our target solution, and we can use inverse optimization to perturb the objective function in each agent's problem to ensure the optimality of the central planner's solution for all the agents. The main tool in the inverse optimization is the resource price function coefficients (i.e., r_i, k). Taking this adjustment into account, the new feasibility problem has far less decision variables, and it is as follows.

$$\text{[P–CP]} \quad \max \quad 0$$

$$
\begin{aligned}
\text{s.t.} \quad & -\pi_j + \sum_i a_{ij} y_{ij} = 0, && \forall j \in J \\
& r_i + 2\, t_{ij}^{c*} \times k - y_{ij} - u_{ij} + v_{ij} = 0, && \forall i \in R, \forall j \in J \\
& a_{ij} x_j^{c*} \leq b_{ij} + t_{ij}^{c*}, && \forall i \in R, \forall j \in J \\
& t_{ij}^{c*} \leq \bar{b}_{ij}, && \forall i \in R, \forall j \in J \\
& -t_{ij}^{c*} \leq b_{ij}, && \forall i \in R, \forall j \in J \\
& x_j^{c*} \geq 0, && \forall j \in J \\
& y_{ij}, u_{ij}, v_{ij} \geq 0, && \forall i \in R, \forall j \in J \\
& y_{ij} \times \left(b_{ij} + t_{ij}^{c*} - a_{ij} x_j^{c*} \right) = 0, && \forall i \in R, \forall j \in J \\
& u_{ij} \times \left(b_{ij} + t_{ij}^{c*} \right) = 0, && \forall i \in R, \forall j \in J \\
& v_{ij} \times \left(\bar{b}_{ij} - t_{ij}^{c*} \right) = 0, && \forall i \in R, \forall j \in J
\end{aligned}
$$

The only decision variables in this problem are the price function coefficients, (r_i, k), and the dual variables, (y_{ij}, u_{ij}, v_{ij}). Since the primal feasibility constraints are already satisfied, we can remove them from the program (P–CP). We denote the central planner's optimal solution by (x^*, t^*), dropping the c superscript for simplicity. The resulting feasibility problem to find the optimal resource price function parameters is the following:

$$
\begin{aligned}
\text{[P–CP]} \quad \max \quad & 0 \\
\text{s.t.} \quad & -\pi_j + \sum_i a_{ij} y_{ij} = 0, && \forall j \in J && (2.71) \\
& r_i + 2\, t_{ij}^{*} \times k - y_{ij} - u_{ij} + v_{ij} = 0, && \forall i \in R, \forall j \in J && (2.72) \\
& y_{ij}, u_{ij}, v_{ij} \geq 0, && \forall i \in R, \forall j \in J && (2.73) \\
& y_{ij} \times \left(b_{ij} + t_{ij}^{*} - a_{ij} x_j^{*} \right) = 0, && \forall i \in R, \forall j \in J && (2.74) \\
& u_{ij} \times \left(b_{ij} + t_{ij}^{*} \right) = 0, && \forall i \in R, \forall j \in J && (2.75) \\
& v_{ij} \times \left(\bar{b}_{ij} - t_{ij}^{*} \right) = 0, && \forall i \in R, \forall j \in J && (2.76)
\end{aligned}
$$

According to the second stationarity condition in (P–CP), equation (2.72), the formula for the constant part of the nonlinear price function, r_i, is as follows:

$$r_i = y_{ij} + u_{ij} - v_{ij} - 2k\, t_{ij}^* \tag{2.77}$$

In order for the price function (2.45) to work, the right–hand–side of (2.77) needs to be the same value for all the agents. Otherwise, coordination on the basis of nonlinear prices without price discrimination will not be possible. Now, we define some new notation. We partition the set of agents and the set of resources into two disjoint subsets, based on the central planner's solution (primal values = x^*, t^*, dual values = p^*):

- **Resources**:
  1. $R^f = \{i : \sum_j a_{ij}x_j^* = b_i\}$, the set of fully–used resources. The entire endowment of this type of resource is used for production purposes. The optimal dual resource price for these types of resources are generally non-zero and positive ($p_i^* > 0$), although it is possible for ($p_i^* = 0$).
  2. $R^p = \{i : \sum_j a_{ij}x_j^* < b_i\}$, the set of partially–used resources. This type of resource is not used to its capacity. The optimal dual resource price for these types of resources are zero ($p_i^* = 0$).
- **Agents:**
  1. J^+ = set of producer agents. These agents have a non–zero production quantity ($x_j^* > 0$). Note that we assume strictly positive a_{ij} coefficients for all the agents; that is, if an agent produces any amount of her product, she needs all resources for production purposes. Producer agents can potentially sell a subset of their resource endowments.
  2. J^0 = set of non–producer agents. These agents make zero production ($x_j^* = 0$), and they rely on selling their resources to benefit from the alliance.

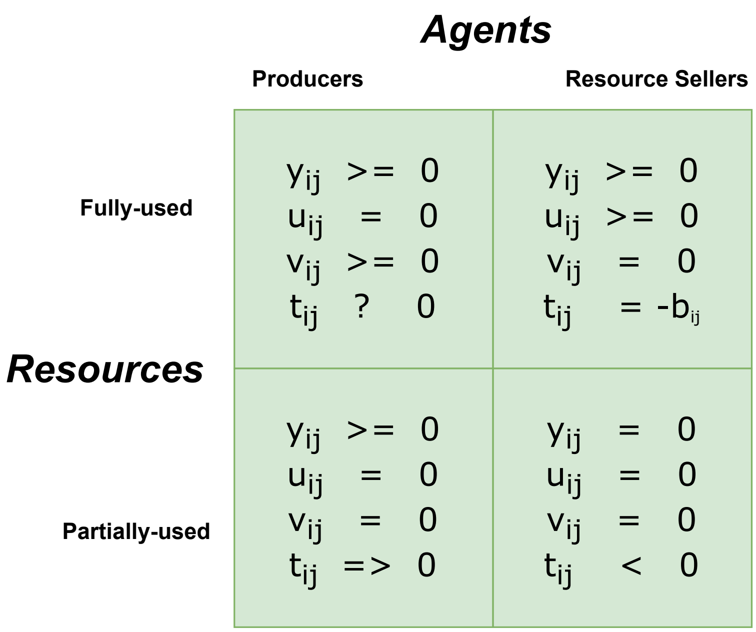


Figure 2.4: Breakdown of complementary slackness conditions for subsets of agents and resources

Using the complementary slackness conditions in (2.74)–(2.76) and the partitioning of the set of agents and resources described above, Figure 2.4 summarizes the breakdown of complementary slackness conditions with respect to each subset of agents and resources.

Taking into account the value of dual variables shown in Figure 2.4, for each subset of the resources (i.e., fully– and partially–used resources), we show that all the conditions in (P–CP), (2.71)–(2.76), can be satisfied.

Partially Used Resources (R^p):

First, let us see what is the resulting equation (2.77) when we commit the assignments in Figure 2.4.

$$r_i = \begin{cases} y_{ij} - 2k\, t^*_{ij}, & \text{if } j \in J^+, \\ -2k\, t^*_{ij}, & \text{if } j \in J^0. \end{cases} \tag{2.78}$$

Now we need to use the two conditions mentioned above, namely minimal exchange and equal division. In essence, we are specifying the desirable exchange values given optimal production quantities. It is not difficult to see that given an optimal production plan (i.e. x^{c*}), there might be multiple optimal exchange values (i.e., t^{c*}), and we have the freedom to choose which pair of (x^{c*}, t^{c*}) we use as the target solution in the inverse optimization approach. Let us denote the average selling quantity by the resource sellers in partially-used case by $\bar{t}_i$, which is equal to:

$$\bar{t}_i = \frac{1}{|J^0|} \sum_{j \in J^+} t^*_{ij} \tag{2.79}$$

We can calculate the exact r_i for the partially-used resources:

$$r_i = 2k\,\bar{t}_i \tag{2.80}$$

Given t_{ij} values are all non-negative for producer agents, we can always satisfy (2.77) using the proper y_{ij}:

$$y_{ij} = r_i + 2k\,t^*_{ij}, \ \forall j \in J^+ \tag{2.81}$$

$$= 2k\,\bar{t}_i + 2k\,t^*_{ij} \ \forall j \in J^+ \tag{2.82}$$

Fully Used Resources (R^f):

The formula for r_i in fully-used resource scenario, taking into account the value assignments in Figure 2.4 is the following:

$$r_i = \begin{cases} y_{ij} - v_{ij} - 2k\,t^*_{ij}, & \text{if } j \in J^+, \\ y_{ij} + u_{ij} + 2k\,b_{ij}, & \text{if } j \in J^0. \end{cases} \tag{2.83}$$

Note that the optimal exchange value for non–producer agents is clear and it is the entire endowment of each agent (i.e., $t^*_{ij} = -b_{ij}$). The optimal exchange variable for producer agents however is not trivial and it depends on their production quantity, thus using the general notation t^*_{ij}. We know that the term $(u_{ij} + 2k\, b_{ij})$ is always non-negative for non–producer agents. The term $(-v_{ij} - 2k\, t^*_{ij})$ may or may not be non–positive depending on the value of t^*_{ij}. It is possible to have a negative (means selling) exchange variable for a producer agent for some resources. Now the goal is to show that:

$$y_{ij_1} - 2k\, t^*_{ij_1} \geq y_{ij_2} + 2k\, b_{ij_2}, \quad \forall i \in R^f, j_1 \in J^+, j_2 \in J^0 \tag{2.84}$$

If we show that (2.84) is always feasible, we can always satisfy conditions (2.83) using proper v_{ij} and u_{ij} dual values. Therefore, the task is to prove that there is always a feasible solution to the following set of constraints:

$$\sum_i a_{ij} y_{ij} = \pi_j, \qquad \forall j \in J \tag{2.85}$$

$$y_{ij_1} - y_{ij_2} \geq 2k \left(t^*_{ij_1} + b_{ij_2} \right), \qquad \forall i \in R^f, j_1 \in J^+, j_2 \in J^0 \tag{2.86}$$

$$y_{ij} \times \left(b_{ij} + t^*_{ij} - a_{ij} x_j \right) = 0, \qquad \forall i \in R, \forall j \in J \tag{2.87}$$

$$y_{ij} \geq 0, \qquad \forall i \in R, \forall j \in J \tag{2.88}$$

Intuitively, for any resource i that is fully used, agents with positive production quantity will have a higher y_{ij} than those agents with zero production. This is particularly clear when we view the coefficients as $\frac{\pi_j}{a_{ij}}$ which is the bang–per–buck for agents.

2.4 Numerical Analysis

To further study and assess the quality of our theoretical arguments, laid out in the previous section, we generate several random instances of the coordinating problem, solve agents and the central planner's problem on each of them, run the settlement process, and calculate the resulting aggregate profit and finally the efficiency ratio. We use a large number of choices for price function parameters (i.e., r, k) in our experiments to better assess the behavior of our target metric, the efficiency ratio, when these parameters are changing. All computational experiments have been written in Python 3.7, and optimization problems are run using Gurobi 9.0 solver.

Table 2.2 contains the information regarding a 3–agent 1–resource example with resource usage and endowment values given, as well as 8 possible profit margin vectors (2 choices for each agent, resulting in 8 different problem instances). We study 8 different profit vectors, to assess the performance of the quadratic pricing function on different objective function values, while keeping a, b parameters constant. These profit margins are numbered $1 \ldots 8$ to facilitate better use of the plots in this section:

- Profit margin #**1**: $(3, 3, 1)$
- Profit margin #**2**: $(3, 3, 2)$
- Profit margin #**3**: $(3, 2, 1)$
- Profit margin #**4**: $(3, 2, 2)$
- Profit margin #**5**: $(5, 3, 1)$
- Profit margin #**6**: $(5, 3, 2)$
- Profit margin #**7**: $(5, 2, 1)$
- Profit margin #**8**: $(5, 2, 2)$

Using a single resource in this example provides additional opportunity to present visualization based on the single base price r in a simpler way. Numerical experiments with larger number of resources have the same message, in general. In this section, we present two main types of visualizations. First, we fix the constant k, and we find the efficiency ratio and the subsidy amount (also known as the central planner's net profit) for each choice of the base price r. Since spot market prices for buying and selling one unit of resource are 4.5 and 1.5, respectively, it is logical to limit our choice of r to be within these two numbers. Otherwise, spot market option will be more attractive than joining and staying within the alliance of

	Problem parameters
Agents	Three agents $J = (j_1, j_2, j_3)$
Resources	One resource $R = (i_1)$
Resource usage	$a = (2, 1, 1)$
Resource endowment	$b = (3, 1, 2)$
Profit margins (π)	$\pi_1 \in \{3, 5\}, \pi_2 \in \{2, 3\}, \pi_3 \in \{1, 2\}$
Base price (r)	Values between 1.6 and 4.4 with increments of 0.01
k values	$k \in \{0, 0.001, 0.01, 0.05, 0.1, 0.2, 0.5, 1, 5, 10, 15, 20\}$
Spot market prices	$p^s = 1.5, p^b = 4.5$

Table 2.2: A 3–agent 1–resource numerical example with 8 possible profit margins

agents. We use approximately 300 r values, ranging from 1.6 to 4.4 with increments of 0.01. We run the same experiment for 8 different profit margin (π) choices, and 12 different k values. Second, for each choice of the constant k, we find the r value with the best (i.e., largest) efficiency ratio, and we plot the resulting efficiency ratio, and profit values (central planner's, aggregate profit, and the centralized profit) as a function of k.

Note that we plot only a subset of all the k values that have been tested (the subset $\{0, 0.001, 0.01, 0.1, 0.5, 1, 10\}$ to be exact) to be able to show the entire instance in one page. The general behavior of the efficiency ratio and the subsidy amount remain the same. According to Figures 2.5–2.12, as k increases, in most instances, the best efficiency ratio falls under the value of one which suggests that there is an optimal point in which the efficiency ratio is at its best. Moving from $k = 0$ to a non–zero k value, the subsidy increases which may impact the effectiveness of nonlinear pricing mechanism negatively. To achieve the best performance, we need to take into account both efficiency ratio and the subsidy at the same time.

For those instances with $k > 0$ where the efficiency ratio is one, especially those closer to zero, there is a range of r values appearing to be the optimal solution. As an important future direction, we need to investigate the behavior of the efficiency ratio within those intervals, and ideally characterize the boundaries of these optimal intervals of r. We also need to take the subsidy into account. Even though there is a range of r values which are optimal in

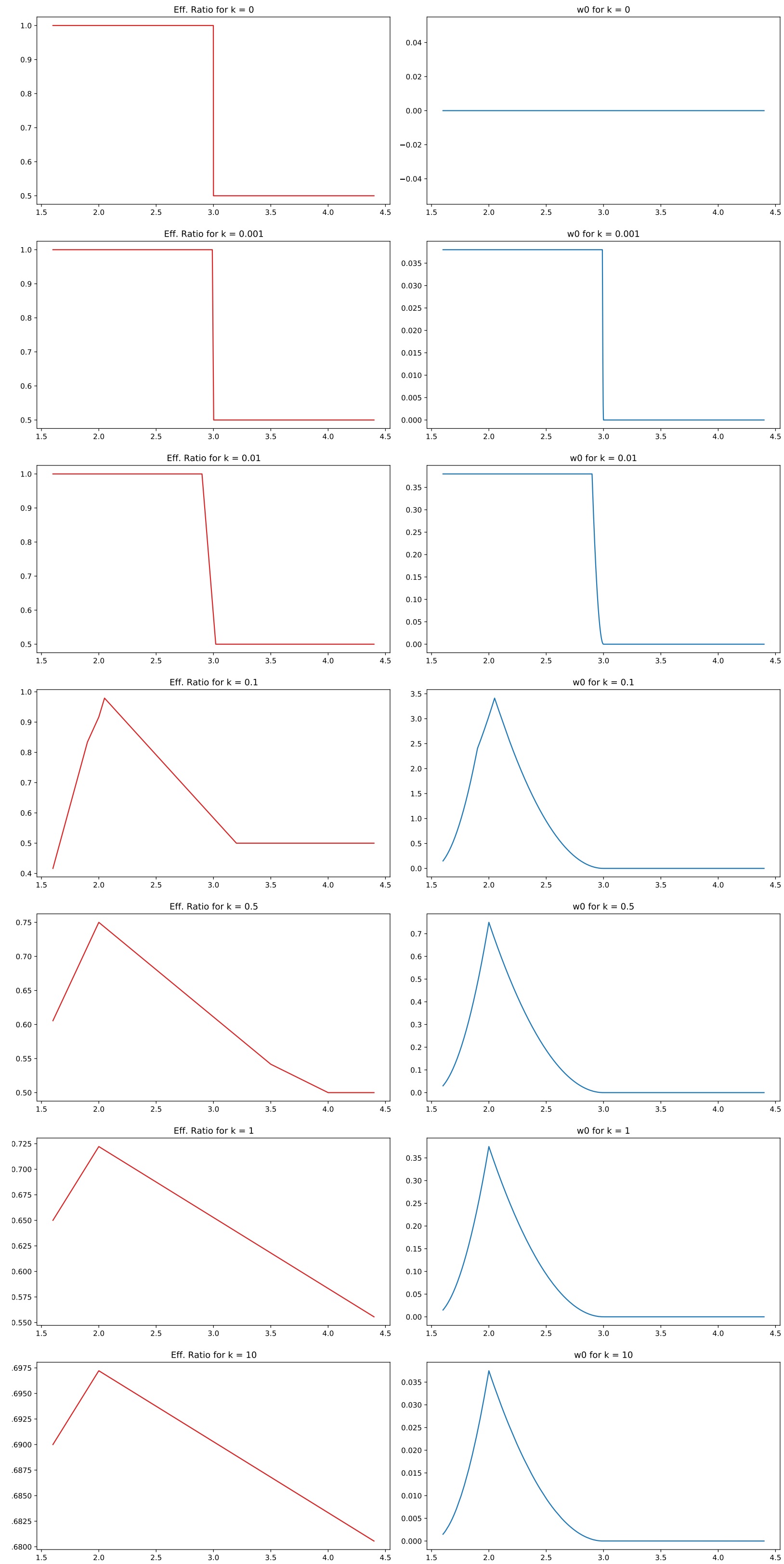


Figure 2.5: Efficiency ratios and subsidy as a function of r for profit margin 1 in example of Table 2.2

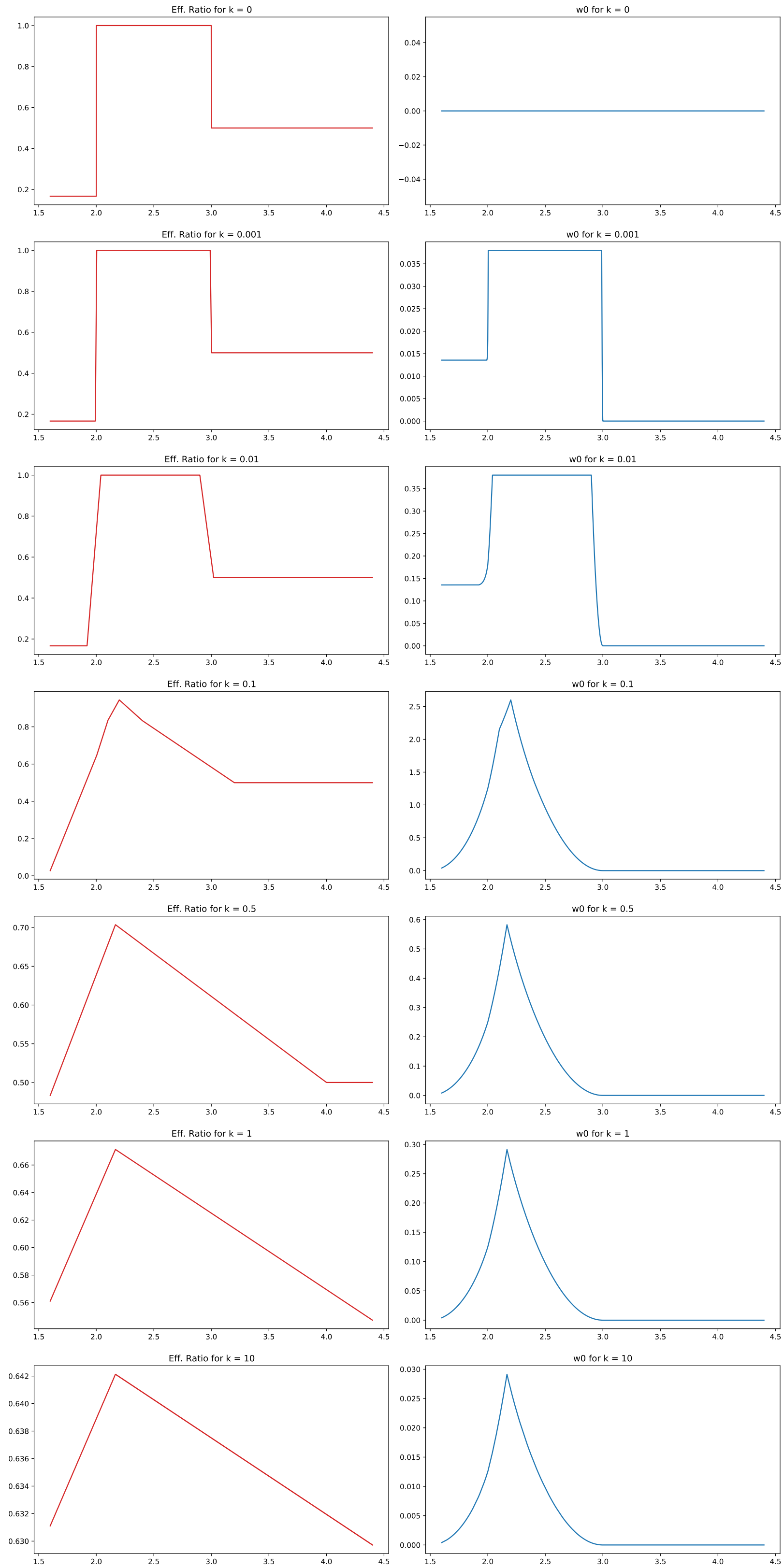


Figure 2.6: Efficiency ratios and subsidy as a function of r for profit margin 2 in example of Table 2.2

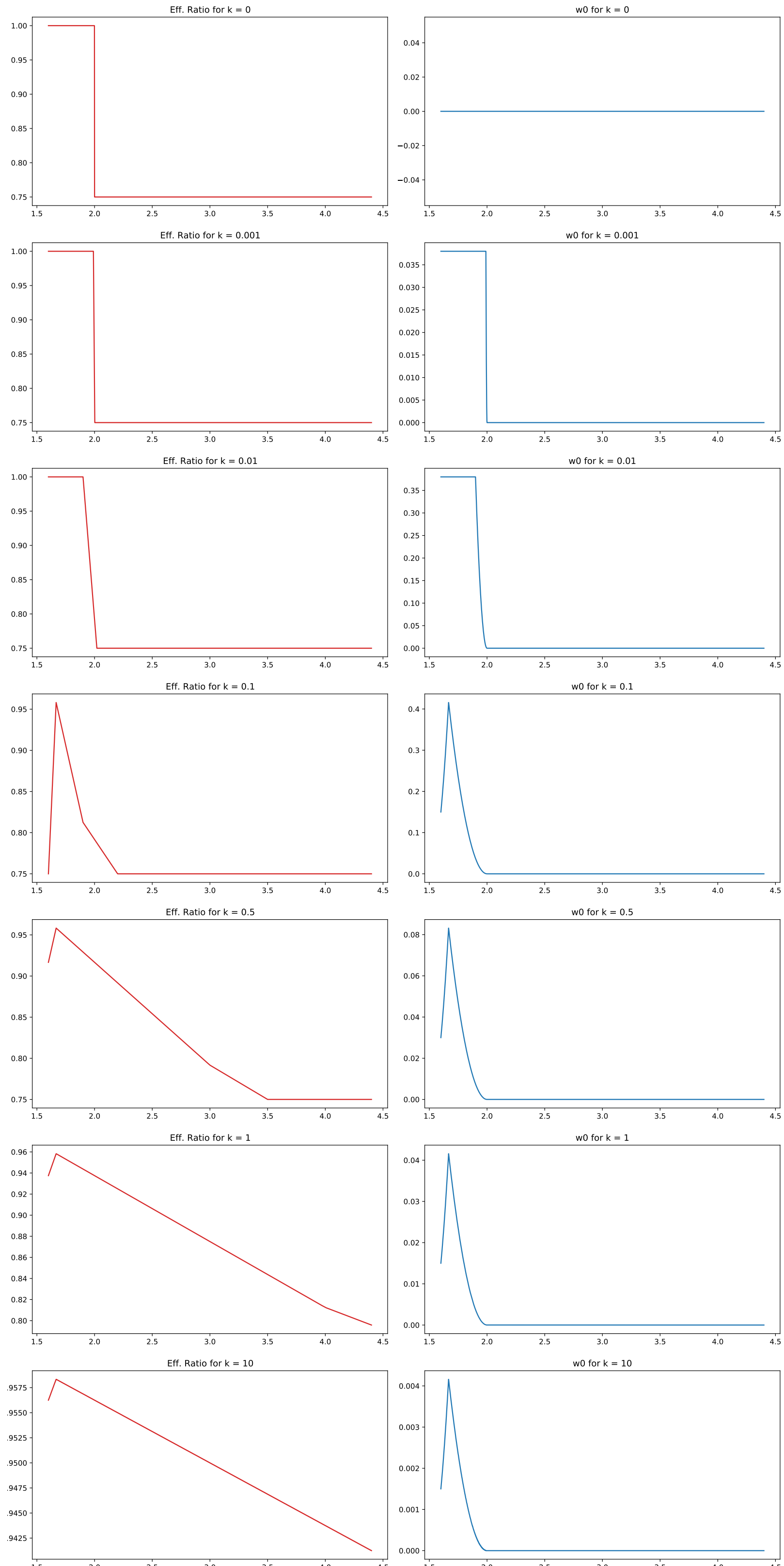


Figure 2.7: Efficiency ratios and subsidy as a function of r for profit margin 3 in example of Table 2.2

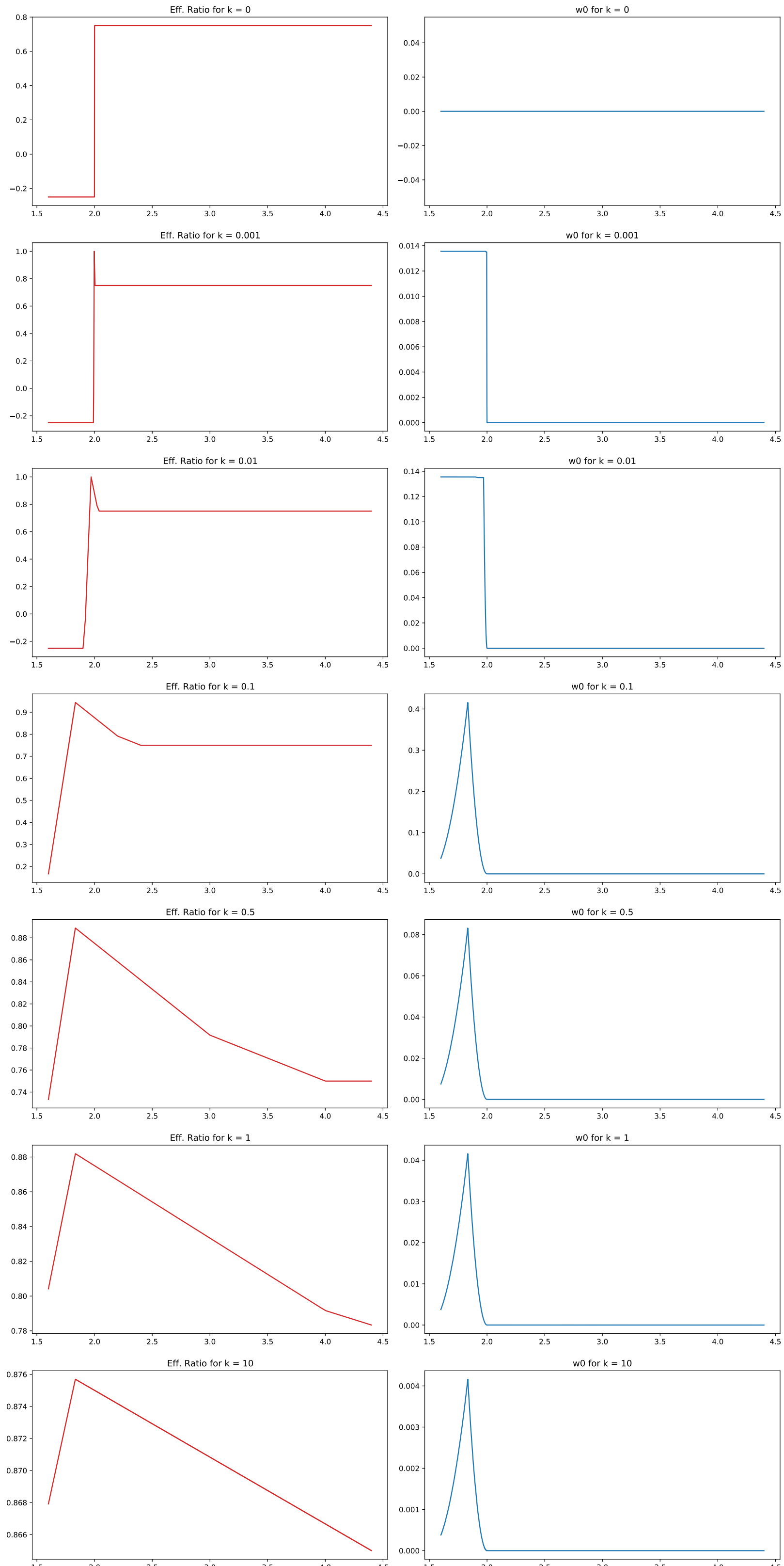


Figure 2.8: Efficiency ratios and subsidy as a function of r for profit margin 4 in example of Table 2.2

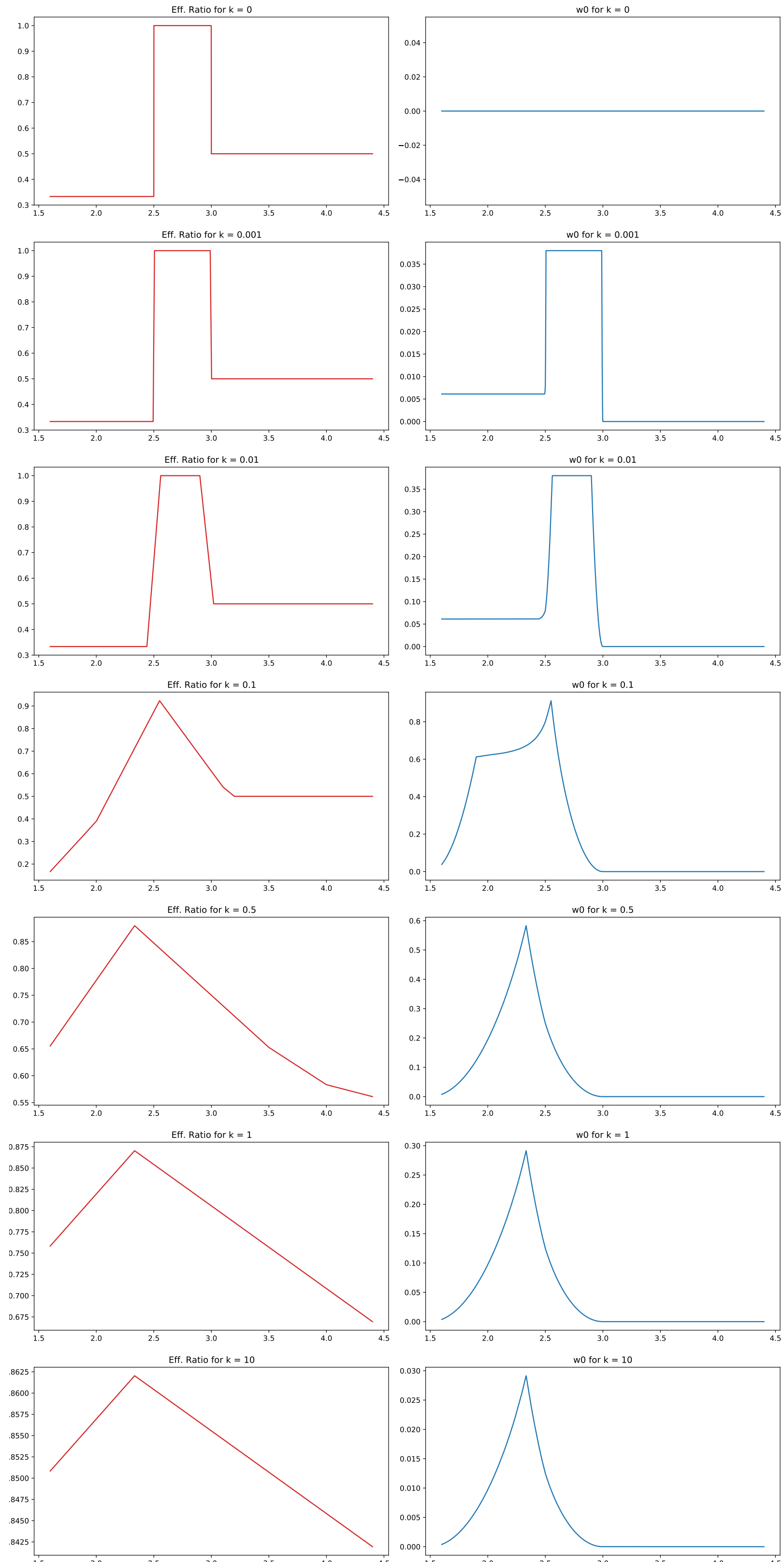


Figure 2.9: Efficiency ratios and subsidy as a function of r for profit margin 5 in example of Table 2.2

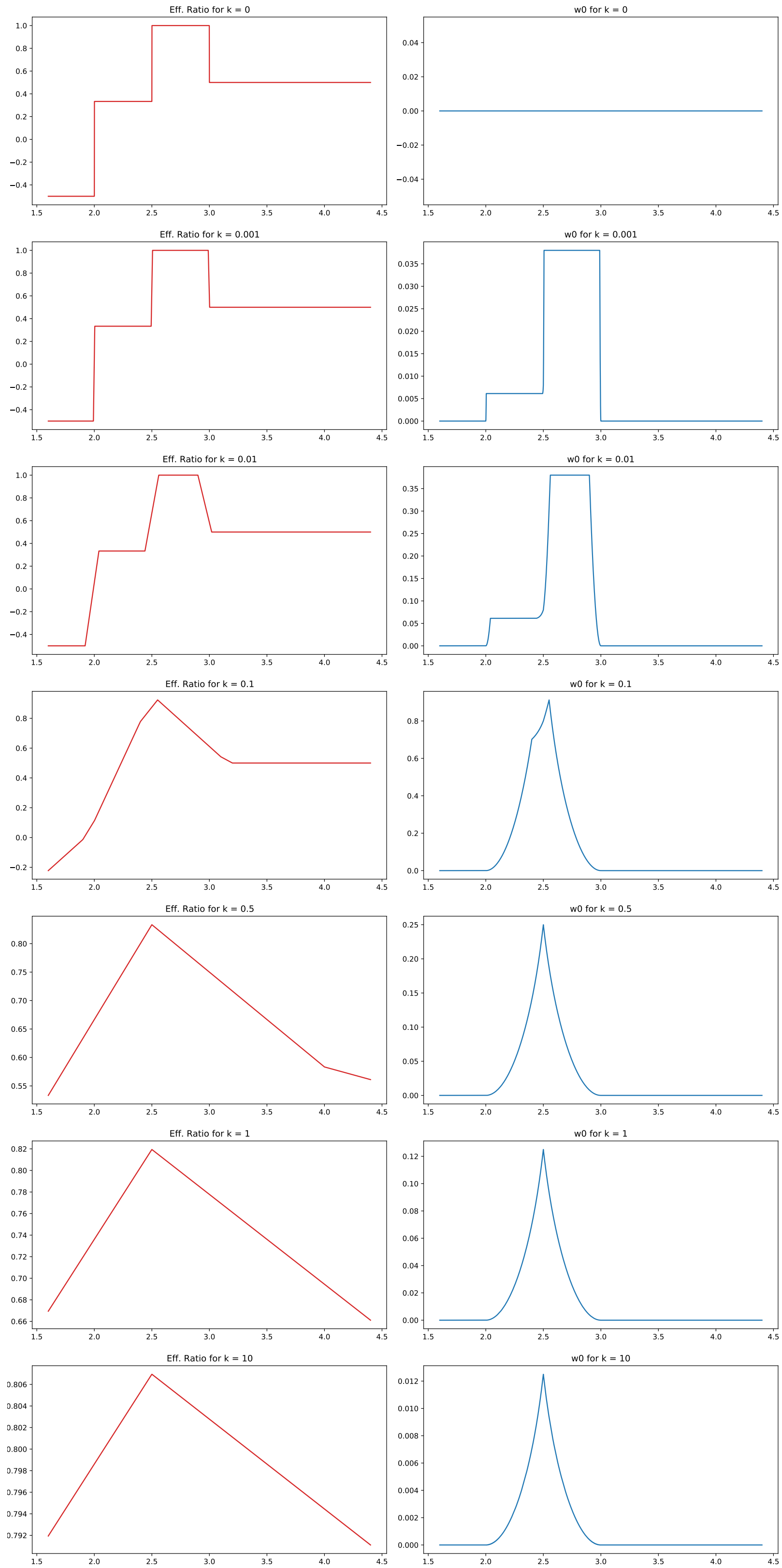


Figure 2.10: Efficiency ratios and subsidy as a function of r for profit margin 6 in example of Table 2.2

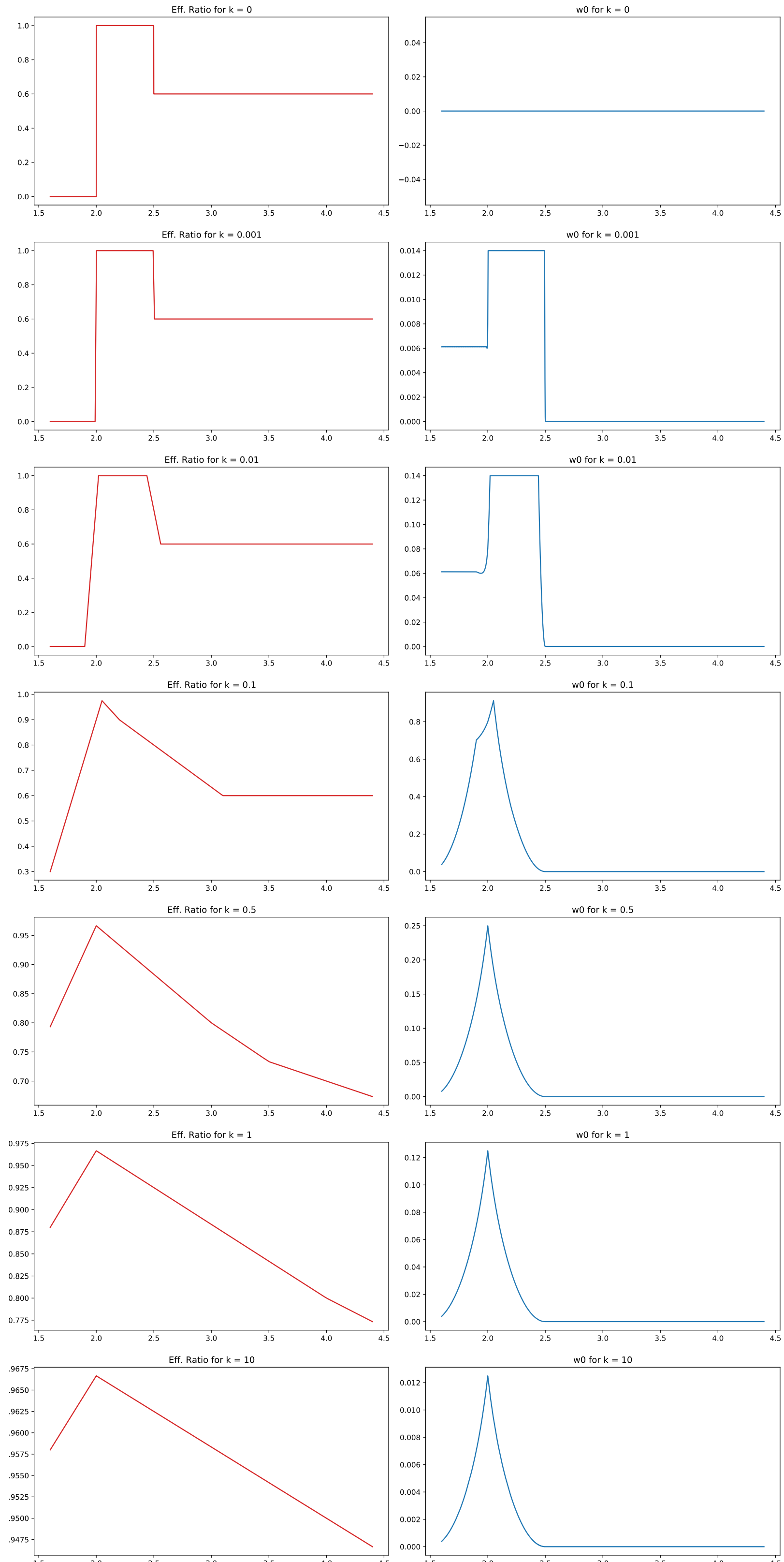


Figure 2.11: Efficiency ratios and subsidy as a function of r for profit margin 7 in example of Table 2.2

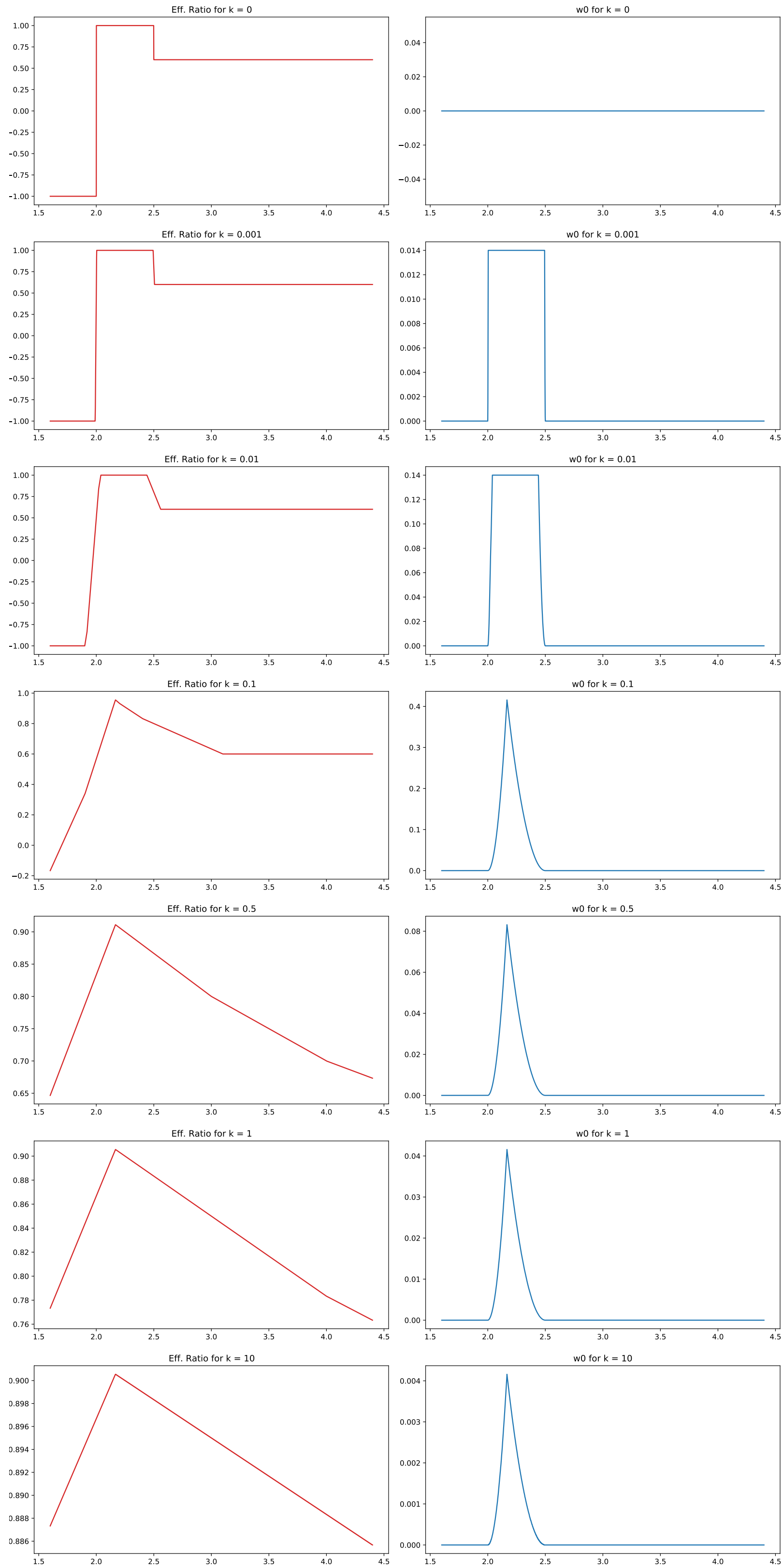


Figure 2.12: Efficiency ratios and subsidy as a function of r for profit margin 8 in example of Table 2.2

terms of efficiency ratio, the subsidy values are not the same for all such points. There can be at least two directions to follow: we can either take a weighted average of these two values (a small example is presented in section 2.5), or we can use the concept of Pareto optimality to consider both metrics at the same time.

Figure 2.13 shows the second experiment we performed to study the effectiveness of nonlinear pricing mechanisms. In these graphs, for each choice of constant k, we find a corresponding r value with the best efficiency ratio, and we pair that efficiency ratio, as well as $(w, w_0, \sum_j w_j)$, with that k value. This figure is a result of plotting these values as functions of k. Figure 2.14 is the same plot, except only the first k values have been plotted to zoom into smaller values of k.

The main message from these graphs is that it helps to use a nonlinear price function to achieve the best efficiency ratio. While increasing the value of k, in all of the graph, the efficiency ratio eventually becomes one at some point, while dropping to a lower value after a certain value for k. This shows the importance of choosing the right k value, while also paying attention to the subsidy amount which can move in the opposite direction.

On the right hand side of these graphs, we plot agents' aggregate profit without the subsidy $(\sum_j w_j)$, the subsidy (w_0), and the overall aggregate profit value (w) as a function of k. One important takeway from this plot is that even though w and w_0 are similar in shape and slope, but we want w to be as large as possible, while w_0 is the opposite.

Finally, we observe that different profit margins have different efficiency ratio values and this variation is a reasonable motivation to pursue this problem in the stochastic case, since combining these instances with a probability distribution, it only gets more complicated to find the most appropriate price function.

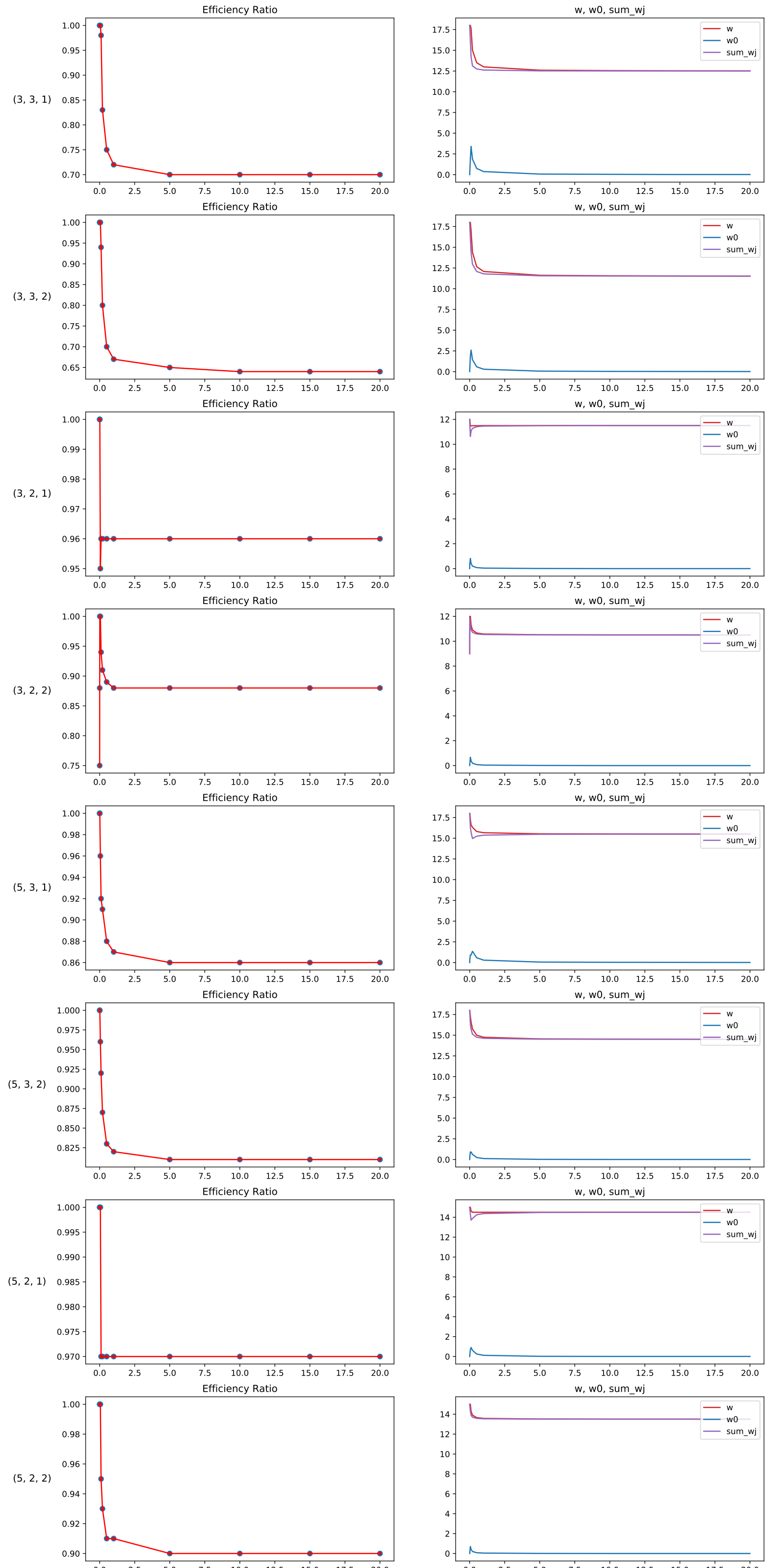


Figure 2.13: Efficiency ratio and profit values as functions of k (each row belongs to a different profit margin)

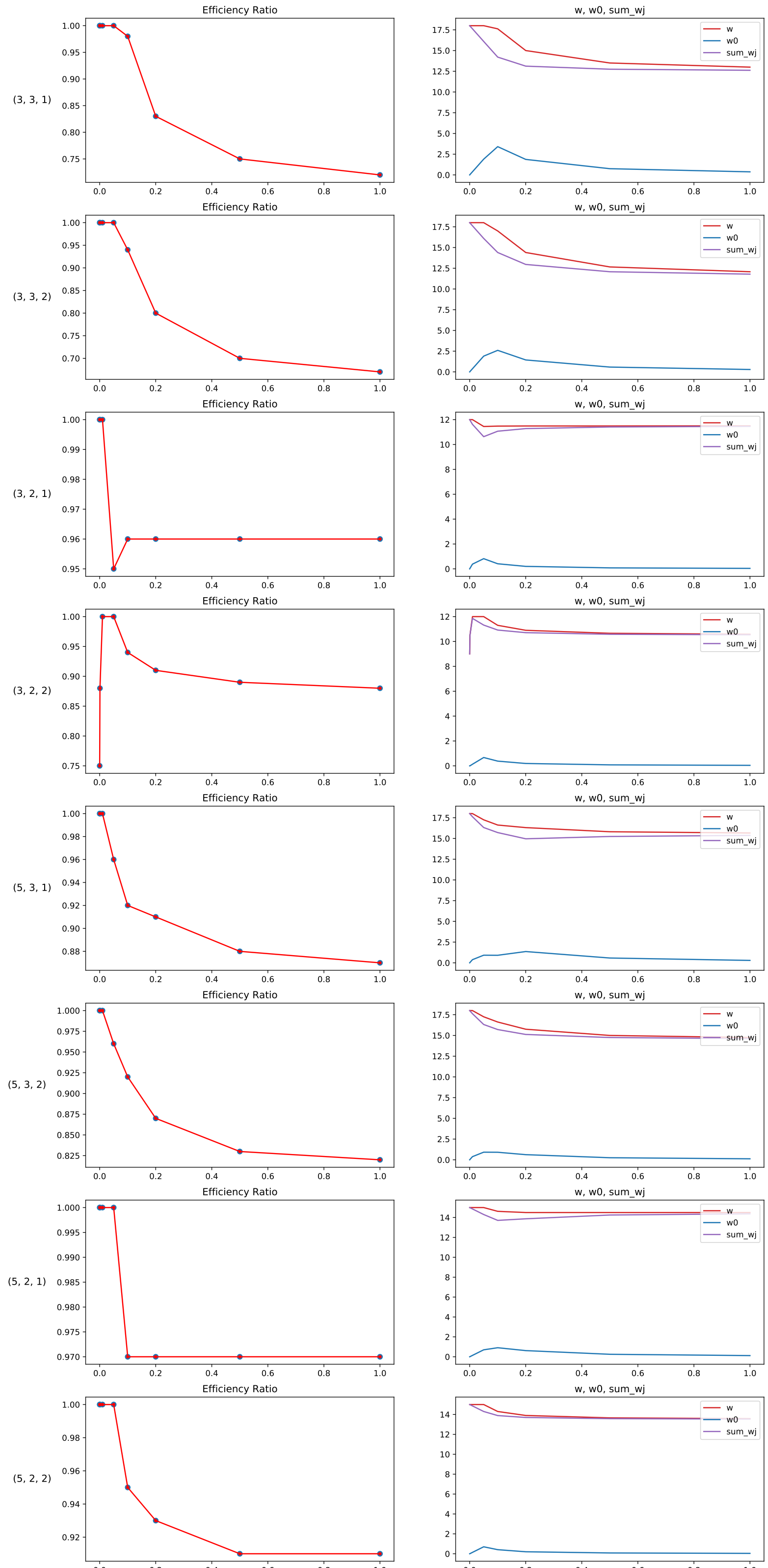


Figure 2.14: Efficiency ratio and profit values as functions of k (first 8 k values)

2.5 Resource Exchange Model - Stochastic Case

In this section, we extend our model by assuming uncertainty in the profit coefficients of the agents. The main implication of this assumption is when the central planner studies the problem parameters to design a pricing schedule, she no longer has full knowledge of the agents' profit coefficients and she only knows a probability distribution of these coefficients. Figure 2.15 illustrates the sequence of events in the stochastic case of the pricing problem in resource exchange economic setting.

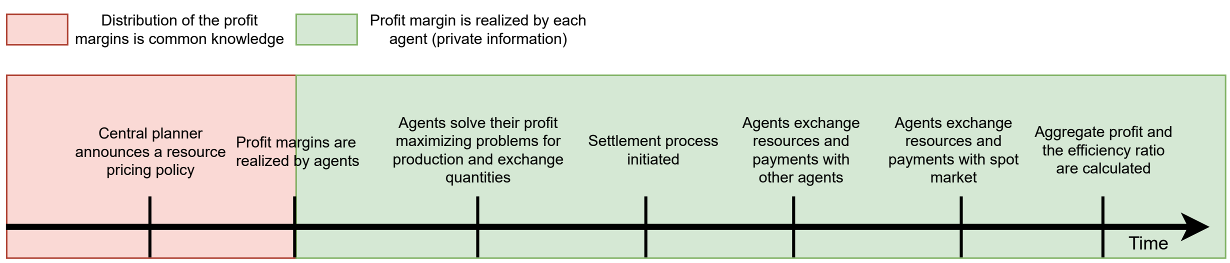


Figure 2.15: Timeline of events in the coordination problem - stochastic case

Let set Ξ denote the set of all profit margin values for agents, and let ξ denote one specific realization of this random variable. We show the vector of profit margins corresponding to the random state ξ by π^ξ, and the specific profit margin for agent j under that scenario by π_j^ξ. For instance, in a two–agent problem where profit margins for agents one and two can both be either 2 or 3, the set Ξ is equal to $\{(2,2),(2,3),(3,2),(3,3)\}$, π^0 refers to the profit vector $(2,2)$, and π_1^0 refers to agent 1's profit margin under this random state, which is 2.

Given the realization ξ of the profit margins, we review agents' profit maximizing problems, denoted by (P_j^s) for agent j, using similar notation as in section 2.3. The price function in (P_j^s), $f_{ij}(t,\Phi)$, is a general price function. In the remainder of this section, we discuss the usefulness of a quadratic pricing policy, similar to the deterministic case, in achieving higher efficiency ratio.

$$[P_j^s]\ z_j^s(\xi) = \max \quad \pi_j^\xi x_j - \sum_i f_{ij}(t,\Phi) \times t_{ij} \tag{2.89}$$

$$\text{s.t.} \quad a_{ij}x_j \leq b_{ij} + t_{ij}, \quad (y_{ij} \geq 0) \qquad \forall i \in R, \forall j \in J \tag{2.90}$$

$$t_{ij} \leq \bar{b}_{ij}, \quad (v_{ij} \geq 0) \qquad \forall i \in R, \forall j \in J \tag{2.91}$$

$$-t_{ij} \leq b_{ij}, \quad (u_{ij} \geq 0) \qquad \forall i \in R, \forall j \in J \tag{2.92}$$

$$x_j \geq 0, \qquad \forall j \in J \tag{2.93}$$

The same process to obtain an approximate coordinating solution in the deterministic case can be applied to find a feasible solution for the stochastic case. Let $w^s(\pi^\xi, \Phi, x, t)$ denote the aggregate profit value given the profit margin π^ξ. The set $\mathbb{H}^\xi$ contains the set of all the optimal solutions (x, t) in agents' individual problems given the profit margin π^ξ.

The central planner's problem on the other hand, which can be used to find the most appropriate resource price function, is modeled below.

$$[\text{M2}^s] \; z_c^s(\xi) = \max \quad \sum_j \pi_j^\xi x_j^c \tag{2.94}$$

$$\text{s.t.} \quad a_{ij}x_j^c \leq b_{ij} + t_{ij}^c, \quad (q_{ij} \geq 0) \qquad \forall i \in R, \forall j \in J \tag{2.95}$$

$$\sum_j t_{ij}^c = 0, \quad (h_i) \qquad \forall i \in R \tag{2.96}$$

$$x_j^c \geq 0, \qquad \forall j \in J \tag{2.97}$$

Let $l(\xi)$ denote the mass density function for the realizations of the profit margins. In our numerical analysis, we assume the set Ξ consists of a countably finite number of scenarios ξ, and each $l(\xi)$ represents the probability of that profit margin, π^ξ, to be the truth. The expected value of the vector of profit margins is calculated as $\mathbb{E}_\xi(\pi) = \sum_{\xi \in \Xi} l(\xi)\pi^\xi$. Given the set of possible profit margins and a probability distribution, we first calculate both the central planner's solution, as well as the decentralized solution according to the process outlined in Figure 2.1 for each choice of the vector π. We then take the expected value of the numerator and the denominator in the efficiency ratio formula, resulting in an efficiency ratio for the stochastic case of the problem.

$$\bar{w}^s(\Phi) = \sum_{\xi \in \Xi} l(\xi) \times \Big(\min_{(x,t) \in \mathbb{H}^\xi} \;\; w^s(\pi^\xi, \Phi, x, t) \Big) \tag{2.98}$$

$$\bar{z}_c^s = \sum_{\xi \in \Xi} l(\xi) z_c^s(\xi) \tag{2.99}$$

Let $\delta^s(\Phi)$ and δ^{s*} denote the efficiency ratio and the best efficiency ratio, respectively, both in the stochastic instance of the problem.

$$\delta^{s*} = \max_{\Phi} \; \delta^s(\Phi) = \max_{\Phi} \; \frac{\bar{w}^s(\Phi)}{\bar{z}_c^s} \tag{2.100}$$

Due to the stochasticity of the problem in this section, studying the effectiveness of different price functions is even more challenging. After performing extensive numerical analysis, we observe the usefulness of quadratic pricing policy for achieving the best efficiency ratio in the stochastic case. Moreover, there is a sweet spot for the parameter k and finding the exact or approximate value of the optimal k remains to be a fruitful direction to explore in the future. Table 2.3 shows the setup for a numerical example with three agents and a single resource. We include a large range of values for the two parameters of the price function (r, k).

	Problem parameters
Agents	Three agents $J = (j_1, j_2, j_3)$
Resources	One resource $R = (i_1)$
Resource usage	$a = (2, 1, 1)$
Resource endowment	$b = (3, 1, 2)$
Profit margins (π)	$\pi_1 \in \{3, 5\}, \pi_2 \in \{2, 3\}, \pi_3 \in \{1, 2\}$
Base price (r)	Values between 1.6 and 4.4 with increments of 0.01
k values	$k \in \{0, 0.001, 0.01, 0.05, 0.1, 0.2, 0.5, 1, 5, 10, 15, 20\}$
Spot market prices	$p^s = 1.5, p^b = 4.5$

Table 2.3: A 3–agent 1–resource numerical example with uncertainty in profit margins

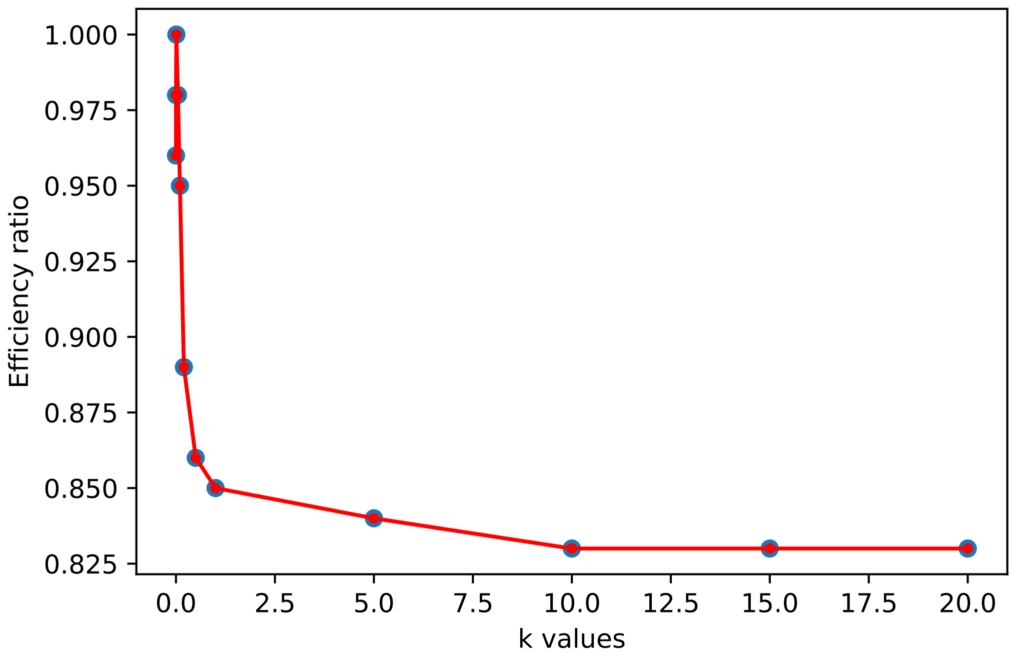


Figure 2.16: Efficiency ratio as a function of k in example of table 2.3

Figure 2.16 illustrates the efficiency ratio values as a function of the constant k, calculated using (2.100). For each choice of the parameter k, approximately 300 different r values have been tested, and the largest efficiency ratio has been selected to pair with that k value in Figure 2.16. Figure 2.17 shows the same plot as in Figure 2.16, except it only contains the first 8 k values, to zoom in the left hand side of the plot, closer to k being zero. Note that the optimal efficiency ratio in this case is not at $(k = 0)$ and not immediately after zero either, as it was the case with nonlinear price function with price discrimination in the deterministic case, and instead, the optimal efficiency ratio happens at a certain point between $k = 0.01$ and $k = 0.05$. We should also mention that, a higher efficiency ratio and the ability to coordinate the agents may come with a price which is the amount of subsidy that the central planner needs to pay the system. Let us plot the subsidy amounts in the same example to further investigate this tradeoff. Figure 2.18 shows the subsidy amount corresponding to the best r value given any choice of k in the x–axis. Figure 2.19 is the same plot as in Figure 2.18, except it only contains the first 8 k values to show the behavior of the plot at the

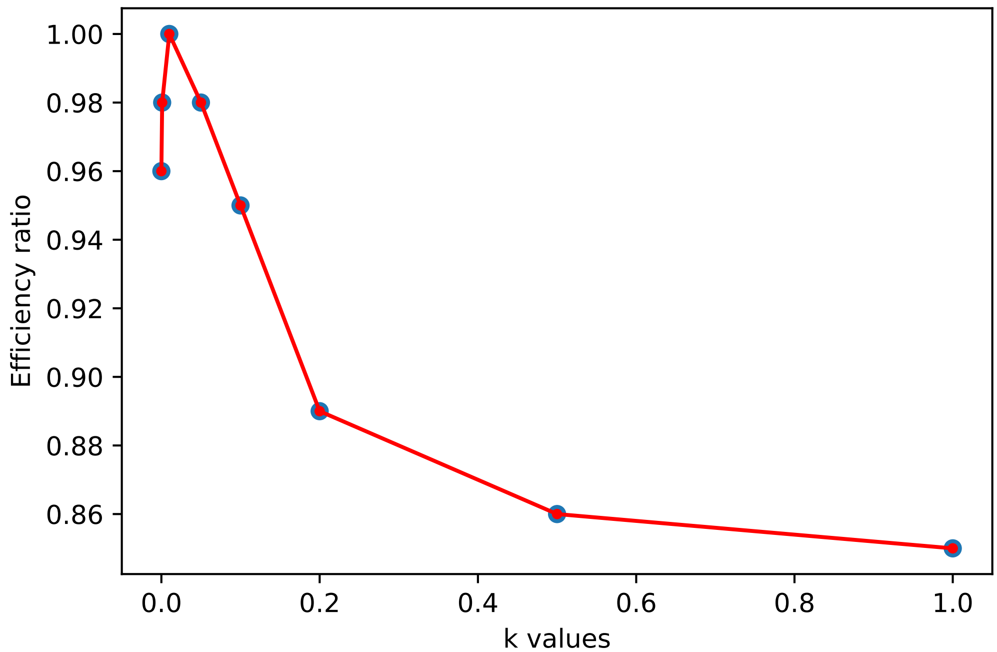


Figure 2.17: Efficiency ratio as a function of k (first 8 k values) in example of table 2.3

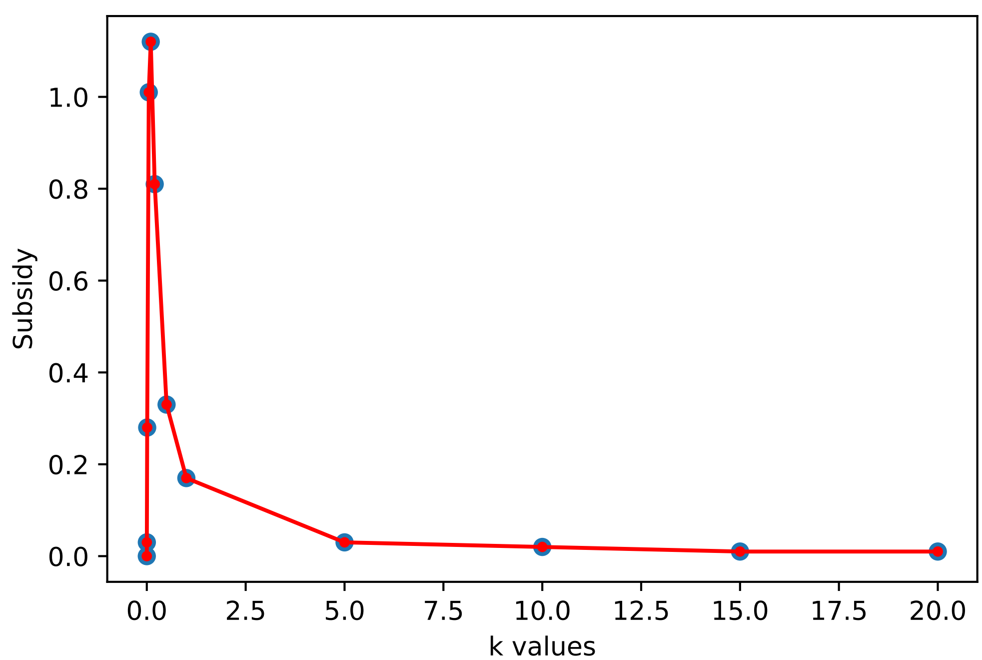


Figure 2.18: Subsidy amount as a function of k in example of table 2.3

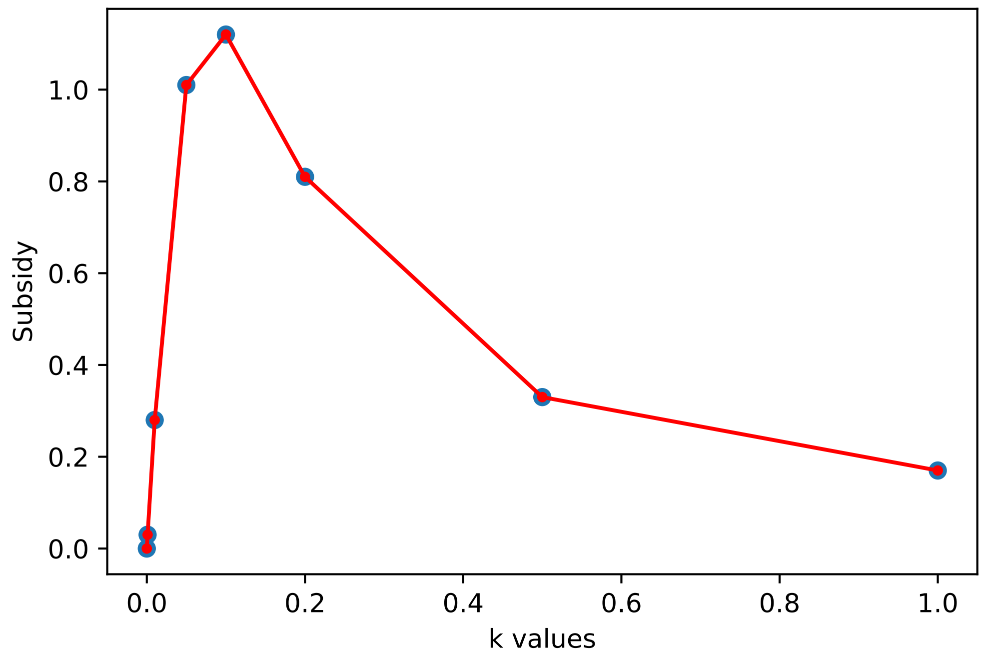


Figure 2.19: Subsidy amount as a function of k (first 8 k values) in example of table 2.3

peak, slightly better. Now, to better compare the advantage of higher efficiency ratio with the drawback of having to subsidize the agents by a large amount, we can use a weighted average of the efficiency ratio and the subsidy amount in a single visualization. Since the efficiecy ratio for a given K ($\delta_k^{s*} = \max_r \delta^s(r,k)$) is a number between zero and one, we convert the subsidy into a percentage, by dividing it by the aggregate profit $\bar{w}_k^{s*} = \max_\Phi \bar{w}^s(\Phi)$. We then subtract this quantity from the efficiency ratio, and we denote it by e_k.

$$e_k = \delta_k^{s*} - \frac{\bar{w}_0^s(k)}{\bar{w}_k^{s*}} \tag{2.101}$$

Figure 2.20 illustrates the tradeoff between the efficiency ratio and the subsidy, converted into the percentage of aggregate profit. Even though the exact choice of the best k value changes from Figure 2.17, but the small range at the lower end of the spectrum for parameter

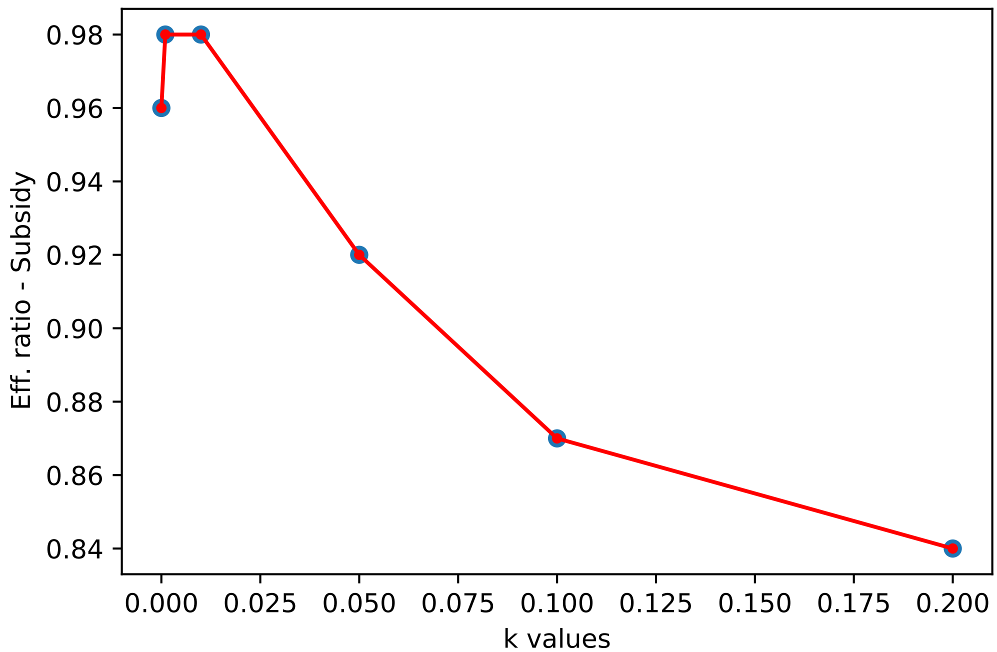


Figure 2.20: Plotting parameter e_k, defined in (2.101), as a function of k (first 6 k values) in example of table 2.3

k remains the best bet. We observed the same behavior in many other numerical examples and different probability distributions, as well. As a future work, we need to investigate the reason why quadratic pricing performs better in achieving descent efficiency ratio in lower k values, and ideally we want to characterize the exact set of k values which maximize δ^{s*}.

2.6 Concluding Remarks

In many real world applications, resources and equipment are being shared in multi–agent economic models for a fee, in an alliance of agents with certain rules and regulations (e.g., the resource price function is announced by the central planner or alliance manager). Each agent has been endowed a certain amount of each resource. These resources and equipment may be used for production purposes, or may be traded with other agents. Agents make two key decisions: how much of their distinct product they are producing, and how much of each type of resource they are trading with other agents. The trade can take both buying and selling form. Agents make decisions oblivious to other agents' decisions. Now the efficiency of the alliance depends on the performance of all agents in making their production and exchange decisions, and whether the overall aggregate solution matches or is close to the central planner's globally optimal solution. We use a metric called efficiency ratio, defined as the ratio between the worst aggregate solution of the decentralized problem and the optimal solution of the centralized problem, to measure the performance of the alliance given any resource price function. We assume that in case of any mismatch between supply and demand of any of the resources, agents can use the spot market to satisfy any unmatched buying or selling request, incurring a higher price, compared to the alliance.

It is well–known in the literature that linear price functions without price discrimination (i.e., same price on each specific resource for all the agents) cannot guarantee an efficiency ratio of one due to the multiple optimal solutions in agents' problems and the possibility of a mismatch between buy/sell orders by the agents. We analyze linear price functions with price discrimination (i.e., linear price for each resource, not necessarily same across different agents) and we argue that it is sufficient to achieve the efficiency ratio of one. The main drawback of this price function however is the fact the central planner needs to subsidize exchanges among agents.

We study nonlinear price functions of the form "constant1 + constant2 * exchange value" and we discuss their advantages and disadvantages. Jennergren has proven the efficiency of this price function with price discrimination. We study a different version of this nonlinear price function, in which the constant part is fixed and the same across agents. We provide the conditions under which this price function achieves the best efficiency ratio, even without price discrimination in the general multiple–resource multiple–agent case.

Finally, we study a different version of this problem where the profit margins of the agents are not fully–known for the central planner when designing the price function. Using numerical analysis, we observe the usefulness of nonlinear price functions of the form studied in the deterministic case of the problem. A key observation in this case is that the second coefficient of the resource price function, k, needs to be at a certain level to maximize the efficiency ratio, and there is both an upper and a lower bound for it, as opposed to the deterministic case where there is only an upper bound, and any sufficiently small value of k solves the coordination problem.

As future work, there are a number of directions that need to be explored. First, one can compute the range of the subsidy amount of the central planner's problem when linear price function with price discrimination is used. Second, a full characterization of the proposed nonlinear price function in the deterministic case may be developed which may include optimality conditions for the parameters of the price function. Third, the instance of the coordination problem with uncertainty in the profit margins of the agents may be further studied either as a stochastic optimization problem or a robust optimization problem.

Chapter 3

Conclusions

In this dissertation, we embarked on a journey to study several key research questions in service operations using prescriptive and predictive analytics. In the first chapter, which is within the general domain of sports analytics, the main challenge was to conclude a sports league after suspension mid–way through the season. We reviewed four possible alternatives to conclude the season, and we proposed a two–phase approach to show why playing a subset (and not all) of the remaining games, which is called a shortened season, can be the fairest compromise. In the first phase, we used a Naïve Bayes binary classification method to predict the outcome of the remaining games. We chose the Naïve Bayes model among five candidate binary classifiers, based on their performance in a training dataset consisting of past NBA seasons, and in terms of a metric which is called predictive power. In the second phase, we investigated three different ways of measuring the distance between a full–season ranking and a shortened–season ranking: concordance, Manhattan distance, and Spearman's ρ coefficient. Using these metrics exactly or inspired by them, we formulated several stochastic optimization models to find the shortened season games. After extensive computational experiments using 14 NBA seasons, in years 2004–2010 and 2012–2018, we reached several important conclusions. First, all five proposed solution methods outperform a baseline greedy algorithm which is non–ranking–based. Second, sample average approximation (SAA) models perform better than mean value approximation (MVP) counterparts, as they include more scenarios of the random variables in their analysis. Third, the deterministic model with quadratic objective function, approximating the exact ranking–based objective, performs better than SAA models both in terms of average concordance and the variation across different seasons. Fourth, after performing statistical hypothesis testing, the two models PC–SAA and PW–DQIP outperform all the other models overall. We outlined a few directions that can potentially improve this work in the future. The predictive model can be enhanced to use data points from previous seasons for training purposes. The formulations in SAA models can be improved to solve larger instances with higher number of scenarios included. PW–DQIP formulation can be improved to solve the problems more efficiently, which translates to less running time and lower optimality gap. One possible direction to explore is using Frank–Wolfe algorithm to approximate the solution in PW–DQIP. In general, analytical studies have been gaining attention in the world of sports, and we hope our work in this chapter is regarded as a good example of how operations research tools can be applied to address major disruptions in sports scheduling.

In the second chapter, we study the effectiveness of linear and nonlinear pricing mechanisms in multi–agent economic settings where agents, endowed with certain resources, exchange their resources amongst each other, while also producing their own unique product. They can make profit by either selling their endowments to other agents, or by producing their product and setting it to outside market. The main metric used in this chapter to measure the efficiency of proposed price functions is efficiency ratio which is the ratio between profit values in the decentralized problem (i.e., aggregate profit from agents' problems after the settlement process) and the centralized problem (i.e., first best solution). Due to the possibility of multiple optimal solutions in agents' problems, a linear resource price function (i.e., a constant unit price on each resource) is not guaranteed to coordinate the agents. We show by numerical analysis that linear pricing with price discrimination (i.e., a different unit price on each resource for different agents) can coordinate the agents. We review nonlinear price functions used in the literature, most of which have price discrimination (i.e., the base price is not identical for all the agents). We prove for the multiple–agent single–resource instance of the problem, that a quadratic price function with a single base price, identical for all agents, can achieve an efficiency ratio of one. We present extensive numerical studies to show that it is highly likely that an efficiency ratio of one can be achieved for the multiple–agent multiple–resource instance of the problem, as well, under certain conditions, and we highlight those conditions. Finally, we study an instance of the of the problem where profit margins of the agents are not fully–known and they are realized only after the resource price function is announced. Using numerical analysis, we show the effectiveness of a quadratic pricing mechanism in practice. This chapter can be improved on a few directions. First, the exact characterization of the nonlinear price function without price discrimination which results in an efficiency ratio of one can be pursued. Second, alternative definitions for the efficiency ratio in the stochastic case can be studied. Third, more complex probability distributions for the profit margins, other than the uniform distribution, can be investigated for the stochastic case of the coordination problem.

Bibliography

Richa Agarwal and ÖZlem Ergun. Mechanism design for a multicommodity flow game in service network alliances. *Operations Research Letters*, 36(5):520–524, 2008.

Ravindra K Ahuja and James B Orlin. Inverse optimization. *Operations Research*, 49(5):771–783, 2001.

Jeremy Arkes and Jose Martinez. Finally, evidence for a momentum effect in the NBA. *Journal of Quantitative Analysis in Sports*, 7(3), 2011.

Kenneth J Arrow. An extension of the basic theorems of classical welfare economics. In *Proceedings of the second Berkeley symposium on mathematical statistics and probability*, pages 507–532. University of California Press, 1951.

Kenneth J Arrow. General economic equilibrium: purpose, analytic techniques, collective choice. *The American Economic Review*, 64(3):253–272, 1974.

Kenneth J Arrow and Gerard Debreu. Existence of an equilibrium for a competitive economy. *Econometrica: Journal of the Econometric Society*, pages 265–290, 1954.

Kenneth J Arrow and Leonid Hurwicz. On the stability of the competitive equilibrium, i. *Econometrica: Journal of the Econometric Society*, pages 522–552, 1958.

Janelle S. Ayres. Surviving COVID-19: A disease tolerance perspective. *Science Advances*, 6(18), 2020.

William J Baumol and Tibor Fabian. Decomposition, pricing for decentralization and external economies. *Management Science*, 11(1):1–32, 1964.

James C Bean and John R Birge. Reducing travelling costs and player fatigue in the national basketball association. *Interfaces*, 10(3):98–102, 1980.

Mehmet A Begen, Retsef Levi, and Maurice Queyranne. A sampling-based approach to appointment scheduling. *Operations Research*, 60(3):675–681, 2012.

Dirk Briskorn and Andreas Drexl. A branch-and-price algorithm for scheduling sport leagues. *Journal of the Operational Research Society*, 60(1):84–93, 2009.

Mark Brown and Joel Sokol. An improved LRMC method for NCAA basketball prediction. *Journal of Quantitative Analysis in Sports*, 6(3), 2010.

Gerard P Cachon. Chapter 6: Supply chain coordination with contracts 1. *Operations Research*, 2003.

So Yeon Chun, Anton J Kleywegt, and Alexander Shapiro. When friends become competitors: The design of resource exchange alliances. *Management Science*, 2016.

Ronald Harry Coase. The nature of the firm. *Economica*, pages 386–405, 1937.

Guido Cocchi et al. Scheduling the Italian national volleyball tournament. *Interfaces*, 48(3):271–284, 2018.

Cohealo. Cohealo, Equipment Sharing for Health Systems. Retrieved on August 1, 2021. `https://cohealo.com/capital-planning/`, August 2021.

George B Dantzig and Philip Wolfe. The decomposition algorithm for linear programs. *Econometrica: Journal of the Econometric Society*, pages 767–778, 1961.

Gerard Debreu. The coefficient of resource utilization. *Econometrica: Journal of the Econometric Society*, pages 273–292, 1951.

Maia Dorsett. Point of no return: COVID-19 and the US health care system. *Science Advances*, 2020.

Andreas Drexl and Sigrid Knust. Sports league scheduling: graph- and resource-based models. *Omega*, 35(5):465–471, 2007.

Guillermo Durán, Mario Guajardo, and Rodrigo Wolf-Yadlin. Operations research techniques for scheduling Chile's second division soccer league. *Interfaces*, 42(3):273–285, 2012.

Kadir Ertogral and S David Wu. Auction-theoretic coordination of production planning in the supply chain. *IIE transactions*, 32(10):931–940, 2000.

ESPN. NBA playoff predictions: Most likely seeds and matchups. Retrieved on May 31, 2020. `espn.com/nba/story/_/id/28830466/nba-playoff-predictions-most-likely-seeds-best-matchups`, Mar 2020.

Mohammad Javad Feizollahi, Mitch Costley, Shabbir Ahmed, and Santiago Grijalva. Large-scale decentralized unit commitment. *International Journal of Electrical Power & Energy Systems*, 73:97–106, 2015.

Charles Fleurent and Jacques A Ferland. Allocating games for the NHL using integer programming. *Operations Research*, 41(4):649–654, 1993.

Nickolas K Freeman, Sharif H Melouk, and John Mittenthal. A scenario-based approach for operating theater scheduling under uncertainty. *Manufacturing & Service Operations Management*, 18(2):245–261, 2016.

Frommer's. A Massive Shortage of Rental Cars Is Confronting U.S. Travelers. Retrieved on August 1, 2021. `https://www.frommers.com/blogs/arthur-frommer-online/blog_posts/a-massive-shortage-of-rental-cars-is-confronting-u-s-travelers`, March 2021.

Michael J Fry and Jeffrey W Ohlmann. Introduction to the special issue on analytics in sports, part I: General sports applications. *Interfaces*, 42(2):105–108, 2012a.

Michael J Fry and Jeffrey W Ohlmann. Introduction to the special issue on analytics in sports, part II: Sports scheduling applications. *Interfaces*, 42(3):229–231, 2012b.

Yong Fu, Mohammad Shahidehpour, and Zuyi Li. Long-term security-constrained unit commitment: hybrid dantzig-wolfe decomposition and subgradient approach. *IEEE Transactions on Power Systems*, 20(4):2093–2106, 2005.

Noah Gans et al. Parametric forecasting and stochastic programming models for call-center workforce scheduling. *Manufacturing & Service Operations Management*, 17(4):571–588, 2015.

Eduardo R Gomes, Quoc Bao Vo, and Ryszard Kowalczyk. Pure exchange markets for resource sharing in federated clouds. *Concurrency and Computation: Practice and Experience*, 24(9):977–991, 2012.

Dries Goossens and Frits Spieksma. Scheduling the Belgian soccer league. *Interfaces*, 39(2):109–118, 2009.

Zhiling Guo, Gary J Koehler, and Andrew B Whinston. A market-based optimization algorithm for distributed systems. *Management Science*, 53(8):1345–1358, 2007.

Martin Henz. Scheduling a major college basketball conference—revisited. *Operations Research*, 49(1):163–168, 2001.

N Hingorani, D Moore, and K Tornqvist. Setting a new course in the container shipping industry. *IBM Business Consulting Services Travel and Transportation*, 2005.

Lori Houghtalen, Özlem Ergun, and Joel Sokol. Designing mechanisms for the management of carrier alliances. *Transportation Science*, 45(4):465–482, 2011.

Justin Hsu, Jamie Morgenstern, Ryan Rogers, Aaron Roth, and Rakesh Vohra. Do prices coordinate markets? In *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*, pages 440–453. ACM, 2016.

Xing Hu, René Caldentey, and Gustavo Vulcano. Revenue sharing in airline alliances. *Management Science*, 59(5):1177–1195, 2013.

Chinwe Peace Igiri and Enoch Okechukwu Nwachukwu. An improved prediction system for football a match result. *IOSR Journal of Engineering (IOSRJEN)*, 4(12):12–20, 2014.

Tiago Januario, Sebastián Urrutia, Celso C Ribeiro, and Dominique De Werra. Edge coloring: A natural model for sports scheduling. *European Journal of Operational Research*, 254(1):1–8, 2016.

Peter Jennergren. Decentralization on the basis of price schedules in linear decomposable resource-allocation problems. *Journal of Financial and Quantitative Analysis*, 7(1):1407–1417, 1972.

Peter Jennergren. A price schedules decomposition algorithm for linear programming problems. *Econometrica: Journal of the Econometric Society*, pages 965–980, 1973.

Joseph Jiaqi Xu, Peter S Fader, and Senthil Veeraraghavan. Designing and evaluating dynamic pricing policies for major league baseball tickets. *Manufacturing & Service Operations Management*, 21(1):121–138, 2019.

Ramesh Johari and John N Tsitsiklis. Network resource allocation and a congestion game. In *Proceedings of the Annual Allerton Conference on Communication Control and Computing*, volume 41, pages 769–778. Citeseer, 2003.

Rinaldo A Jose, Patrick T Harker, and Lyle H Ungar. Coordinating locally constrained agents using augmented pricing. Technical report, Wharton Financial Institutions Center Working Paper, 1997.

Edward H Kaplan. OM Forum—COVID-19 scratch models to support local decisions. *Manufacturing & Service Operations Management*, 2020.

Graham Kendall, Sigrid Knust, Celso C Ribeiro, and Sebastián Urrutia. Scheduling in sports: An annotated bibliography. *Computers & Operations Research*, 37(1):1–19, 2010.

Maurice G Kendall. A new measure of rank correlation. *Biometrika*, 30(1/2):81–93, 1938.

Anton J Kleywegt, Alexander Shapiro, and Tito Homem-de Mello. The sample average approximation method for stochastic discrete optimization. *SIAM Journal on Optimization*, 12(2): 479–502, 2002.

Kent J Kostuk and Keith A Willoughby. A decision support system for scheduling the Canadian football league. *Interfaces*, 42(3):286–295, 2012.

Rhyd Lewis and J Thompson. On the application of graph colouring techniques in round-robin sports scheduling. *Computers & Operations Research*, 38(1):190–204, 2011.

Bernard Loeffelholz, Earl Bednar, and Kenneth W Bauer. Predicting NBA games using neural networks. *Journal of Quantitative Analysis in Sports*, 5(1), 2009.

Mengshi Lu, Zhihao Chen, and Siqian Shen. Optimizing the profitability and quality of service in carshare systems under demand uncertainty. *Manufacturing & Service Operations Management*, 20(2):162–180, 2018.

Rhonda Magel and Yana Melnykov. Examining influential factors and predicting outcomes in European soccer games. *International Journal of Sports Science*, 4(3):91–96, 2014.

S Maheswari and C Vijayalakshmi. A lagrangian decomposition model for unit commitment problem. *International Journal of Computer Applications*, 43(12):21–25, 2012.

Patrick Maillé and Bruno Tuffin. Analysis of price competition in a slotted resource allocation game. In *INFOCOM 2008. The 27th Conference on Computer Communications. IEEE*, pages 888–896. IEEE, 2008.

Alexey Malakhov and Rakesh V Vohra. An optimal auction for capacity constrained bidders: a network perspective. *Economic Theory*, 39(1):113–128, 2009.

NBA. NBA Articles. Retrieved on May 31, 2020. `nba.com/article/2020/03/11/coronavirus-pandemic-causes-nba-suspend-season`, May 2020a.

NBA. NBA Articles. Retrieved on May 31, 2020. `nba.com/article/2020/05/23/nba-talking-disney-about-resuming-season-ap`, May 2020b.

NBA. NBA Statistics Homepage. Retrieved on May 31, 2020. `stats.nba.com`, May 2020c.

George L Nemhauser and Michael A Trick. Scheduling a major college basketball conference. *Operations Research*, 46(1):1–8, 1998.

Guillermo Owen. On the core of linear production games. *Mathematical programming*, 9(1):358–370, 1975.

F. Pedregosa et al. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

Georgia Perakis and Guillaume Roels. The price of anarchy in supply chains: Quantifying the efficiency of price-only contracts. *Management Science*, 53(8):1249–1268, 2007.

Rasmus V Rasmussen and Michael A Trick. Round robin scheduling–a survey. *European Journal of Operational Research*, 188(3):617–636, 2008.

Guillaume Roels and Christopher S Tang. Win-win capacity allocation contracts in coproduction and codistribution alliances. *Management Science*, 2016.

Donald G Saari and Carl P Simon. Effective price mechanisms. *Econometrica: Journal of the Econometric Society*, pages 1097–1125, 1978.

Beth Jones Sanborn. Case study: How kaiser permanente saved $8.6 million sharing medical equipment and reimagined asset management. *Healthcare Finance*, 22, 2018.

Charles Spearman. The proof and measurement of association between two things. *The American Journal of Psychology*, 15(1):72–101, 1904.

Sports Illustrated. NBA power rankings: Looking five years into the future. Retrieved on May 31, 2020. `si.com/nba/2020/04/24/nba-power-rankings-predicting-every-teams-future`, Apr 2020.

Fadi Thabtah, Li Zhang, and Neda Abdelhamid. NBA game result prediction using feature analysis and machine learning. *Annals of Data Science*, 6(1):103–116, 2019.

Alaa Tharwat. Classification assessment methods. *Applied Computing and Informatics*, 2018.

Michael A Trick, Hakan Yildiz, and Tallys Yunes. Scheduling major league baseball umpires and the traveling umpire problem. *Interfaces*, 42(3):232–244, 2012.

Timothy Van Zandt. Firms, prices and markets, 2012.

Holger Voos. Resource allocation in continuous production using market-based multi-agent systems. In *Industrial Informatics, 2007 5th IEEE International Conference on*, volume 2, pages 1085–1090. IEEE, 2007.

Léon Walras. Elements of pure economics; or, the theory of social wealth, 1969.

Howard J Weiss. The bias of schedules and playoff systems in professional sports. *Management Science*, 32(6):696–713, 1986.

Stephan Westphal. Scheduling the German basketball league. *Interfaces*, 44(5):498–508, 2014.

WHO. Coronavirus disease (COVID-19) pandemic. Retrieved on June 22, 2021. `who.int/emergencies/diseases/novel-coronavirus-2019`, June 2021.

Mike B Wright. Scheduling fixtures for basketball New Zealand. *Computers & Operations Research*, 33(7):1875–1893, 2006.

Appendix A

Supplement to Concluding a Suspended Sports League

A.1 Proof of Theorems and Propositions

Proposition 1.1. *Predictive power as defined in* (1.6) *measures the expected accuracy provided that the outcome of game* $g \in G$*, denoted* W_g*, follows a Bernoulli distribution with probability* p_g*.*

Proof. Proof. We need to compute $\mathbb{E}[\mathbb{I}(W_g = y_g])$, where $\mathbb{I}(\cdot)$ is the indicator function. Since W_g is Bernoulli, $\mathbb{E}[\mathbb{I}(W_g = y_g]) = P(W_g = y_g)$. For $y_g = 1$, $P(W_g = y_g) = p_g$, and for $y_g = 0$, $P(W_g = y_g) = 1 - p_g$. It follows that for arbitrary y_g we have $\mathbb{E}[\mathbb{I}(W_g = y_g)] = y_g \times p_g + (1 - y_g) \times (1 - p_g)$. □

Proposition 1.2. *For any shortened season* $x \in X$ *and realization* $\xi \in \Xi$*, let* $\varphi_C(x, \xi)$ *and*

$\varphi_M(x,\xi)$ be the objective values of PC and PM, respectively. The following relationship holds:

$$\varphi_M(x,\xi) \leq (n-1) - \varphi_C(x,\xi) \tag{A.1}$$

Proof. Proof. Using the definition of concordant pairs per team we have

$$\begin{aligned}\varphi_C(x,\xi) =& \frac{1}{n}\sum_{i\in T}\sum_{j\in T:j\neq i}(z_{ij}(x,\xi)\hat{z}_{ij}(\xi) + (1-z_{ij}(x,\xi))(1-\hat{z}_{ij}(\xi)))\\ =& \frac{1}{n}\sum_{i\in T}\sum_{j\in T:j\neq i}(1-|z_{ij}(x,\xi)-\hat{z}_{ij}(\xi)|) = (n-1) - \frac{1}{n}\sum_{i\in T}\sum_{j\in T:j\neq i}|z_{ij}(x,\xi)-\hat{z}_{ij}(\xi)|.\end{aligned}$$

On the other hand, by definition of $r(x,\xi)$ and $\hat{r}(\xi)$ as given in (1.17) we have

$$\begin{aligned}\varphi_M(x,\xi) = \frac{1}{n}\sum_{i\in T}|r_i(x,\xi)-\hat{r}_i(\xi)| &= \frac{1}{n}\sum_{i\in T}\left|\sum_{j\in T:j\neq i}(z_{ij}(x,\xi)-\hat{z}_{ij}(\xi))\right|\\ &\leq \frac{1}{n}\sum_{i\in T}\sum_{j\in T:j\neq i}|z_{ij}(x,\xi)-\hat{z}_{ij}(\xi)|,\end{aligned}$$

where the inequality holds by triangle inequality. □

Theorem 1.1. *The stochastic model PW can be solved using the following equivalent deterministic linearly constrained quadratic optimization problem*

$$\text{[PW-DQIP]}\ \min \quad \sum_{i\in T}\left(\frac{1}{m^2}(v_i+\mu_i^2) + \frac{1}{\hat{m}^2}(\hat{v}_i+\hat{\mu}_i^2) - \frac{2}{m\hat{m}}(v_i+\mu_i\hat{\mu}_i)\right) \tag{A.2}$$

$$s.t.\quad \mu_i = y_i^0 + \sum_{g\in G_i^h} p_g x_g + \sum_{g\in G_i^a}(1-p_g)x_g \tag{A.3}$$

$$v_i = \sum_{g\in G_i^h\cup G_i^a} p_g(1-p_g)x_g \tag{A.4}$$

$$x \in X, \tag{A.5}$$

where the decision variables, in addition to $x = \{x_g, g \in G\}$, include μ_i and v_i which encode the mean and variance of the number of wins for team i in the shortened season, respectively. Moreover, the following parameters represent the mean and variance of the number of wins for team i in the full season, respectively:

$$\begin{aligned}
\hat{\mu}_i = \mathbb{E}_\xi\left[\hat{y}_i(\xi)\right] =& y_i^0 + \sum_{g \in G_i^h} p_g + \sum_{g \in G_i^a} (1 - p_g) \\
\hat{v}_i = \mathbb{V}_\xi\left[\hat{y}_i(\xi)\right] =& \sum_{g \in G_i^h \cup G_i^a} p_g(1 - p_g).
\end{aligned}$$

Proof. Proof. We may expand the expectation in the objective function (1.26) as

$$\begin{aligned}
&\mathbb{E}_\xi\left[\sum_{i \in T}\left(\frac{y_i(\xi)}{m} - \frac{\hat{y}_i(\xi)}{\hat{m}}\right)^2\right] = \sum_{i \in T} \mathbb{E}_\xi\left[\left(\frac{y_i^2(\xi)}{m^2} + \frac{\hat{y}_i^2(\xi)}{\hat{m}^2} - 2\frac{y_i(\xi)}{m}\frac{\hat{y}_i(\xi)}{\hat{m}}\right)\right] \\
=& \sum_{i \in T}\left(\frac{1}{m^2}\mathbb{E}_\xi\left[y_i^2(\xi)\right] + \frac{1}{\hat{m}^2}\mathbb{E}_\xi\left[\hat{y}_i^2(\xi)\right] - \frac{2}{m\hat{m}}\mathbb{E}_\xi\left[y_i(\xi)\hat{y}_i(\xi)\right]\right).
\end{aligned}$$

In the following, we compute the individual expectation terms.

1. To compute $\mathbb{E}_\xi\left[y_i^2(\xi)\right]$, we restate $\mathbb{E}_\xi\left[y_i^2(\xi)\right]$ as $\mathbb{E}_\xi\left[y_i^2(\xi)\right] = \mathbb{V}_\xi\left[y_i(\xi)\right] + \mathbb{E}_\xi\left[y_i(\xi)\right]^2$, where $\mathbb{V}(\cdot) = Var(\cdot)$ stands for variance. The expected number of wins for team i in the shortened season (i.e., $\mathbb{E}_\xi\left[y_i(\xi)\right]$) is

$$\begin{aligned}
\mathbb{E}_\xi\left[y_i(\xi)\right] =& \mathbb{E}_\xi\left[y_i^0 + \sum_{g \in G_i^h} W_g(\xi)x_g + \sum_{g \in G_i^a}(1 - W_g(\xi))x_g\right] \\
=& y_i^0 + \sum_{g \in G_i^h} p_g x_g + \sum_{g \in G_i^a}(1 - p_g)x_g
\end{aligned}$$

Since the outcome of games are assumed to be independent, we may compute the variance of the number of wins for team i in the shortened season (i.e., $\mathbb{V}_\xi\left[y_i(\xi)\right]$) as

$$\mathbb{V}_\xi\left[y_i(\xi)\right] = \sum_{g\in G_i^h} \mathbb{V}_\xi\left[W_g(\xi)x_g\right] + \sum_{g\in G_i^a} \mathbb{V}_\xi\left[(1-W_g(\xi))x_g\right].$$

Since $W_g(\xi)$ is Bernoulli, and x_g is binary, we have

$$\mathbb{V}_\xi\left[W_g(\xi)x_g\right] = \mathbb{V}_\xi\left[(1-W_g(\xi))x_g\right] = x_g^2 p_g(1-p_g) = x_g p_g(1-p_g),$$

yielding

$$\mathbb{V}_\xi\left[y_i(\xi)\right] = \sum_{g\in G_i^h\cup G_i^a} p_g(1-p_g)x_g.$$

2. By the same token, we may compute $\mathbb{E}_\xi\left[\hat{y}_i^2(\xi)\right]$ using the identity $\mathbb{E}_\xi\left[\hat{y}_i^2(\xi)\right] = \mathbb{V}_\xi\left[\hat{y}_i(\xi)\right] + \mathbb{E}_\xi\left[\hat{y}_i(\xi)\right]^2$, where

$$\begin{aligned}\mathbb{E}_\xi\left[\hat{y}_i(\xi)\right] =& \mathbb{E}_\xi\left[y_i^0 + \sum_{g\in G_i^h} W_g(\xi) + \sum_{g\in G_i^a}(1-W_g(\xi))\right] = y_i^0 + \sum_{g\in G_i^h} p_g + \sum_{g\in G_i^a}(1-p_g),\\ \mathbb{V}_\xi\left[\hat{y}_i(\xi)\right] =& \sum_{g\in G_i^h} \mathbb{V}_\xi\left[W_g(\xi)\right] + \sum_{g\in G_i^a} \mathbb{V}_\xi\left[(1-W_g(\xi))\right] = \sum_{g\in G_i^h\cup G_i^a} p_g(1-p_g).\end{aligned}$$

3. To compute $\mathbb{E}_\xi\left[y_i(\xi)\hat{y}_i(\xi)\right]$, we first expand the expression as

$$\mathbb{E}_\xi\left[y_i(\xi)\hat{y}_i(\xi)\right] = y_i^0\mathbb{E}_\xi\left[\hat{y}_i(\xi)\right] + \sum_{g\in G_i^h} x_g\mathbb{E}_\xi\left[W_g(\xi)\hat{y}_i(\xi)\right] + \sum_{g\in G_i^a} x_g\mathbb{E}_\xi\left[(1-W_g(\xi))\hat{y}_i(\xi)\right].$$

For $g \in G_i^h$ we have

$$\mathbb{E}_\xi\left[W_g(\xi)\hat{y}_i(\xi)\right] = y_i^0\mathbb{E}_\xi\left[W_g(\xi)\right] + \sum_{q\in G_i^h} \mathbb{E}_\xi\left[W_g(\xi)W_q(\xi)\right] + \sum_{q\in G_i^a} \mathbb{E}_\xi\left[W_g(\xi)(1-W_q(\xi))\right]$$

$$=p_g \left(y_i^0 + 1 + \sum_{q \in G_i^h : q \neq g} p_q + \sum_{q \in G_i^a} (1-p_q) \right) = p_g \left(1 - p_g + \mathbb{E}_\xi \left[\hat{y}_i(\xi)\right]\right),$$

where we have used $\mathbb{E}_\xi \left[W_g^2(\xi)\right] = p_g(1-p_g) + p_g^2 = p_g$. Similarly, for $g \in G_i^a$ we have

$$\begin{aligned}
\mathbb{E}_\xi \left[(1 - W_g(\xi))\hat{y}_i(\xi)\right] =& \mathbb{E}_\xi \left[\hat{y}_i(\xi)\right] - \mathbb{E}_\xi \left[W_g(\xi)\hat{y}_i(\xi)\right] \\
=& \mathbb{E}_\xi \left[\hat{y}_i(\xi)\right] - p_g \left(y_i^0 + \sum_{q \in G_i^h} p_q + \sum_{q \in G_i^a : q \neq g} (1-p_q) \right) \\
=& \mathbb{E}_\xi \left[\hat{y}_i(\xi)\right] - p_g \left(-(1-p_g) + \mathbb{E}_\xi \left[\hat{y}_i(\xi)\right]\right) \\
=& p_g(1-p_g) + (1-p_g)\mathbb{E}_\xi \left[\hat{y}_i(\xi)\right],
\end{aligned}$$

where we have used $\mathbb{E}_\xi \left[W_g(\xi)(1 - W_g(\xi))\right] = 0$. Consequently, we may state $\mathbb{E}_\xi \left[y_i(\xi)\hat{y}_i(\xi)\right]$ as

$$\begin{aligned}
\mathbb{E}_\xi \left[y_i(\xi)\hat{y}_i(\xi)\right] =& \mathbb{E}_\xi \left[\hat{y}_i(\xi)\right] \left(y_i^0 + \sum_{g \in G_i^h} x_g p_g + \sum_{g \in G_i^a} x_g (1-p_g) \right) + \sum_{g \in G_i^h \cup G_i^a} x_g p_g (1-p_g) \\
=& \mathbb{E}_\xi \left[y_i(\xi)\right] \mathbb{E}_\xi \left[\hat{y}_i(\xi)\right] + \mathbb{V}_\xi \left[y_i(\xi)\right]
\end{aligned}$$

Putting these pieces together we may restate the objective function (1.26) as

$$\mathbb{E}_\xi \left[\sum_{i \in T} \left(\frac{y_i(\xi)}{m} - \frac{\hat{y}_i(\xi)}{\hat{m}} \right)^2 \right] = \sum_{i \in T} \left(\frac{1}{m^2}(v_i + \mu_i^2) + \frac{1}{\hat{m}^2}(\hat{v}_i + \hat{\mu}_i^2) - \frac{2}{m\hat{m}}(v_i + \mu_i \hat{\mu}_i) \right).$$

$\square$

A.2 Naïve Bayes Classifier

In a Naïve Bayes classifier, the discrimination rule is characterized by a conditional probability of the form $p(y|\vec{x})$ to determine the class label $y \in \{0, 1\}$ for a given observation $\vec{x}$.

Applying Bayes theorem, we may state $p(y|\vec{x})$ as

$$p(y|\vec{x}) = \frac{p(y) \times p(\vec{x}|y)}{p(\vec{x})},$$

in which $p(y)$ is the prior probability of class y (i.e., proportion of data points in the training set with class label y), $p(\vec{x}|y)$ is the likelihood of observing $\vec{x}$ in data points with class label y, and $p(\vec{x}) = p(0)p(\vec{x}|0) + p(1)p(\vec{x}|1)$ is a normalization scalar. Using the naïve conditional independence assumption, we have $p(\vec{x}|y) = \prod_{j=1}^{D} p(x_j|y)$ yielding

$$p(y|\vec{x}) = \frac{1}{p(\vec{x})} \left(p(y) \times \prod_{j=1}^{D} p(x_j|y) \right).$$

Figure A.1 presents the distribution of features conditioned on the class labels across different data points. As illustrated in this figure, the likelihood probability distribution of features conditioned on the class labels approximately follow Gaussian distributions. Therefore, we use Gaussian distribution to represent the likelihood of features conditioned on the binary target variable (i.e., $p(x_j|y) \sim \mathcal{N}(\mu_{j,y},\, \sigma^2_{j,y})$). The Gaussian probability density function for feature x_j conditioned on class label y is as follows

$$p(x_j|y) = \frac{1}{\sigma_{j,y}\sqrt{2\pi}} \exp\left(\frac{-\left(x_j - \mu_{j,y}\right)^2}{2\sigma^2_{j,y}} \right) \tag{A.6}$$

The parameters $\mu_{j,y}$ and $\sigma_{j,y}$ in (A.6) defined for each feature x_j and class label y are estimated using maximum likelihood, and the estimated values for the 2019–20 NBA season are given in Table A.1.

Figure A.1: Scatterplot Matrix of all the eight features in the predictive model

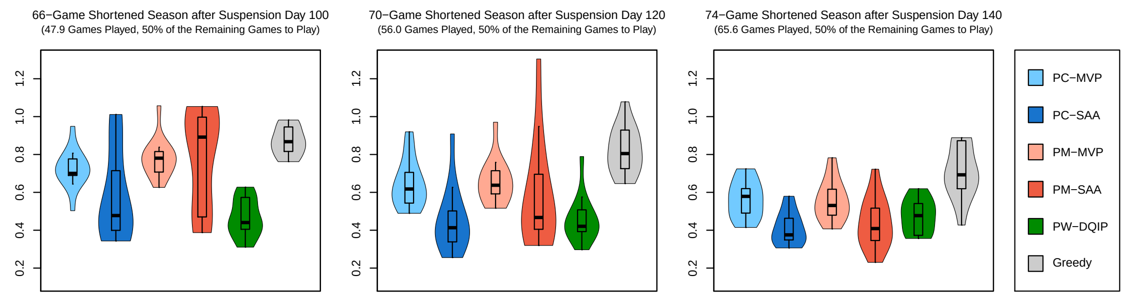
A.3 Supplementary Results

A.3.1 Monte Carlo Simulation Results

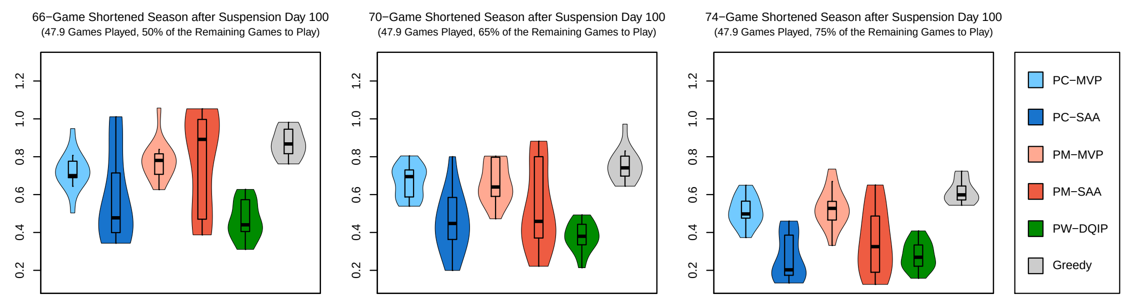
Figure A.2 presents the simulation results according to Manhattan distance per team (lower is better) across different choices of suspension day and target number of games per team over 14 seasons. Figure A.3 shows the results of t-tests using Manhattan distance per team. These results are in agreement with the simulation and t-test results according to concordance per team as presented in Figures 1.8 and 1.9.

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
$\mu_{j,0}$	0.43	-1.93	0.47	0.47	0.56	1.70	0.55	0.51
$\mu_{j,1}$	0.55	1.47	0.53	0.62	0.46	-1.28	0.47	0.40
$\sigma_{j,0}$	0.02	20.9	0.02	0.03	0.03	25.4	0.03	0.03
$\sigma_{j,1}$	0.03	24.3	0.03	0.04	0.02	21.3	0.02	0.02

Table A.1: Estimated parameters using equation (A.6) in the NBA season 2019–20

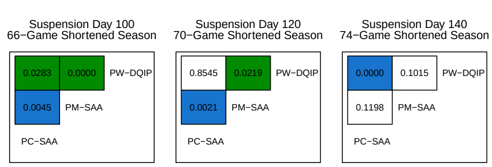


(a) Target fixed at 50% of the remaining games (i.e., shortened season with 66, 70, and 74 games)

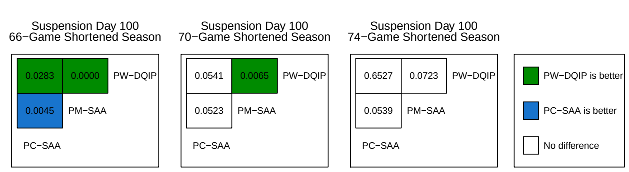


(b) Suspension day fixed at 100

Figure A.2: Distribution of the simulation results (Manhattan distance per team) across 14 NBA seasons



(a) Instances corresponding to Figure A.2(a)



(b) Instances corresponding to Figure A.2(b)

Figure A.3: Results of paired t-tests (Manhattan distance) and the p-value for our three top performing models