2016

Three essays on multi-level optimization models and applications

Mohammad Rahdar
Iowa State University

Three essays on multi-level optimization models and applications

by

Mohammad Rahdar

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Industrial Engineering

Program of Study Committee:

Lizhi Wang, Co-major Professor

Guiping Hu, Co-major Professor

Caroline Krejci

Sarah M Ryan

Zhijun Wu

Iowa State University

Ames, Iowa

2016

DEDICATION

I would like to dedicate this dissertation to the memory of my late father, who I have not seen him for more than twenty years, but never forgotten. To my mother for her ongoing love and support and to my brother for his words of inspiration and encouragement.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to express the deepest appreciation to my advisors Dr. Lizhi Wang and Dr. Guiping Hu, who are tremendous mentors, supportive friends, and thoughtful leaders, for their detailed guidance, patience, understanding, and for providing me with a pleasant atmosphere for doing research. Besides my advisors, I would like to thank the rest of my dissertation committee: Dr. Caroline Krejci, Dr. Sarah M. Ryan, and Dr. Zhijun Wu for their constructive advice, insightful comments, and encouragement. Without their valuable support, it would not be possible to conduct this research.

I would also like to express my gratitude to my friends and colleagues for making my life an enjoyable experience at Iowa State University.

Last but not the least, I would like to thank my family: my mother and my brother for devoting themselves to my life and career, and for their genuine and unconditional love and support.

ABSTRACT

The general form of a multi-level mathematical programming problem is a set of nested optimization problems, in which each level controls a series of decision variables independently. However, the value of decision variables may also impact the objective function of other levels. A two-level model is called a bilevel model and can be considered as a Stackelberg game with a leader and a follower. The leader anticipates the response of the follower and optimizes its objective function, and then the follower reacts to the leader's action.

The multi-level decision-making model has many real-world applications such as government decisions, energy policies, market economy, network design, etc. However, there is a lack of capable algorithms to solve medium and large scale these types of problems. The dissertation is devoted to both theoretical research and applications of multi-level mathematical programming models, which consists of three parts, each in a paper format.

The first part studies the renewable energy portfolio under two major renewable energy policies. The potential competition for biomass for the growth of the renewable energy portfolio in the United States and other interactions between two policies over the next twenty years are investigated. This problem mainly has two levels of decision makers: the government/policy makers and biofuel producers/electricity generators/farmers. We focus on the lower-level problem to predict the amount of capacity expansions, fuel production, and power generation. In the second part, we address uncertainty over demand and lead time in a multi-stage mathematical programming problem. We propose a two-stage tri-level optimization model in the concept of rolling horizon approach to reducing the dimensionality of the multi-stage problem. In the third part of the dissertation, we introduce a new branch and bound algorithm to solve bilevel linear programming problems. The total time is reduced by solving a smaller relaxation problem in each node and decreasing the number of iterations. Computational experiments show that the proposed algorithm is faster than the existing ones.

CHAPTER 1. GENERAL INTRODUCTION

A multi-level mathematical programming problem consists of a hierarchical decision structure with conflicting or compatible objectives. If there are only two levels of decision making, and objective functions and constraints are linear, it is called a Bilevel Linear Programming (BLP) problem. Moreover, it can be considered as a Stackelberg game, which is a strategic game with a leader and a follower. In this game, the leader moves first, and then the follower reacts rationally to the leader's action. Since it is assumed that the information about objective functions and constraints is fully shared with both levels [Bard and Moore (1990)], the leader anticipates the response of the follower and optimizes its objective function; after that, the follower reacts rationally to optimize its objective by considering the action of the upper-level decision maker. Mathematically, bilevel programming problems have two sets of variables, pertaining to upper and lower-level decision makers, respectively. The optimization problem of the follower is enclosed within the constraints of the leader problem. Since there is an optimization problem within the constraints of the upper-level problem, the solution of the upper-level problem is feasible only if this solution is optimal to the lower-level problem.

The Multi-level decision-making model has many useful real-world applications such as government decisions, energy policies, market economy, transportation, supply chain, network design, etc. Bard et al. (2000) proposed a bilevel programming approach to determine tax credits of biofuel production. In their paper, the government was the leader and would like to establish the level of tax credits in biofuel industry such that the annual tax credits would be minimized. The agricultural sector was the follower and wanted to maximize its profits by defining the land used for nonfood crops. Lu et al. (2006) studied a real case of a road network problem to improve it. The leader was the road management committee, and they would like to minimize the system travel cost. The followers were public and private traffic user groups, which sought to minimize their travel delays uncooperatively. Dempe et al. (2005) presented a bilevel

programming model to minimize the cash-out penalties of a natural gas shipper. The leader was a shipper who delivered natural gas from a receipt to a delivery meter, and the follower was the pipeline. Since there always exist operational and transportation imbalances (the difference between the amount of nominated and actually transported) in transporting natural gas, pipelines issue penalties for higher imbalances. The shipper would like to maximize their revenue and pipelines wanted to minimize the amount of cash transaction. In addition, Chiou (2005) used a bilevel programming technique to determine the link capacity expansions and the equilibrium flows in a continuous network design problem. Saranwong and Likasiri (2016) found the best locations for distribution centers by using a bilevel programming approach, where the upper-level problem determined the optimal locations of distribution centers, and the lower-level problem assigned each distribution center to customers to satisfy demands. Moreover, Camacho-Vallejo et al. (2015) proposed a bilevel mathematical model to optimize humanitarian logistics when a catastrophic disaster happened. Kuo and Han (2011) used bilevel linear programming as a decentralized decision modeling in a supply chain distribution system where the leader and the follower were distribution centers and manufacturers, respectively. Tookanlou et al. (2015) proposed a bilevel model to determine the hourly energy prices for electricity. The upper-level problem determined the energy prices to minimize the annual costs of Combined Cooling, Heating, and Power (CCHP) systems. The lower-level problem was distribution utility, which sought to maximize its annual income by selecting an amount of purchased electricity from wholesale electricity market and CCHP system.

In addition to bi-level programming models, several studies have been done on developing and solving tri-level optimization models. Jin and Ryan (2014) proposed a tri-level model of generation and transmission expansion planning problem. Centralized decisions on transmission expansion were made in the first level, decisions of multiple decentralized power generation companies on generation expansion were made in the second level, and multiple market players' operational decisions were made in the third level. Chen et al. (2014) minimized the maximum cost and maximum regret of the transmission expansion planning problem to obtain a robust solution under uncertainty by developing a tri-level mixed integer model.

The dissertation is devoted to both theoretical research and applications of multi-level mathematical programming models, which consists of three parts. In the first part, we studied the renewable energy portfolio under two major renewable energy policies. Since there is a hierarchical relation between decision makers, it is a multi-level decision-making problem. Upper-level decision makers are the government and policy makers, and the lower-level decision makers are producers and farmers. In this study, we focused on the lower-level problem given the policies and decisions of the upper-level decision makers. In the second part, we addressed the uncertainty of demand and lead time in a supply chain by developing a tri-level inventory control optimization model. Uncertain demand and lead time are observed in each period; thus, it is a multi-stage decision-making problem. We approximated the multi-stage problem as a two-stage problem by developing a tri-level model and using the rolling horizon approach. In the third part of the dissertation, we studied a theoretical aspect of bilevel linear programming problems by focusing on the design of a new branch and bound algorithm, considering that most current algorithms cannot solve medium and large scale bi-level programming problems. Our primary results show that the new algorithm is faster than branch and bound method, which is one of the most efficient algorithm to solve this type of problems. The research is introduced in more detail in the remainder of this chapter.

In the first paper, we reduced a complicated multi-level decision-making problem to a simple problem for prediction purposes. It focuses on the potential competition for biomass from RPS[1] and RFS2[2] as two major policies for the growth of the renewable energy portfolio in the United States over the next twenty years. Since a full understanding of the short-term outcome and long-term implications of such competition is demanding by policy makers and other stakeholders of the renewable energy industry, we were motivated to study the interactions between two major renewable policies, RPS and RFS2, and the influence of these interactions on the growth of the U.S. renewable energy portfolio over the next two decades. A huge amount of research has been done on these two major policies separately, with limitations on resource and geographical dimensions. Most focused on the subset of the renewable energy portfolio

[1]Renewable Portfolio Standard
[2]Renewable Fuel Standard

within a geographical region, and few have examined the interactions between RPS and RFS2 or the implications of such interactions on the comprehensive renewable energy portfolio.

This problem has multiple levels of decision making including the government and policy makers, biofuel producers, electricity generators, and farmers. Developing such a model is almost impractical, and it is exceedingly difficult to solve. However, there are mainly two levels of decision makers: the government/policy makers and biofuel producers/electricity generators/farmers; and we focused on the lower-level problem to predict the amount of capacity expansions, fuel production, and power generation. The decision variables of the upper-level problem are the wholesale electricity price, tax credits, RPS/RFS2 requirements, and penalties for non-compliance with RPS/RFS2 policies, which are considered as determined parameters in the model to solve the lower-level problem. Furthermore, we performed a sensitivity analysis to study how the nationwide total renewable energy generation would be adjusted if these assumed known parameters were changed. The proposed model can be extended to multilevel programming problems for decision-making purposes, such as determining the energy price, tax credits, RPS and RFS2 requirements and penalties.

For the second paper, we proposed a tri-level inventory control optimization model to find a robust solution when demands and lead-times are uncertain. A supply chain's efficiency and effectiveness depend on the organization's ability to understand and manage supply and demand uncertainties. However, this has been proven to be a major challenge. In particular, when manufacturers have insufficient information to accurately predict downstream demands and upstream lead times, it is very difficult for them to determine the order sizes and reorder points that will minimize total cost. Since uncertain demand is observed in each period and the exact lead time is not realized until whenever the order arrives, the lack of perfect information about demand and lead time expands the problem to a multi-stage decision-making problem. To approximate the multi-stage decision-making problem and reduce its dimensionality, we proposed a two-stage tri-level optimization model, in which the first period of the planning horizon is the first stage of the problem and all the remaining periods are aggregated into the second stage. Therefore, after making decisions in the first stage, all uncertain parameters are assumed to be observable; thus, the second stage becomes a deterministic problem. As a result,

the two-stage decision-making model can be formulated as a tri-level optimization model. The upper-level makes the first stage decision, the middle-level defines the worst case scenario given the first stage decision, and the lower-level makes the second stage decision considering both the first stage decision and the worst case scenario. This simplified formulation was run in a rolling horizon framework (Beaudin and Zareipour, 2015), in which the model is solved in every period with updated information but only the first stage decisions are implemented. To solve the tri-level model, we designed and implemented an effective algorithm.

In the third paper, we focused on designing an efficient algorithm to solve bilevel linear programming problems, which are NP-hard. There is a lack of capable algorithms to solve medium and large-scale bi-level programming problems; thus, we were motivated to work on developing new algorithms or improving current methods to solve these types of problems more efficiently. Sakawa and Nishizaki (2009) argued that bilevel programming problems are non-convex problems, even if the objective functions and constraints of both levels are linear. This makes bilevel optimization problems difficult to solve. Bard (1991) discussed the difficulties of developing efficient algorithms to solve the BLP problems and proved that the BLP problem is NP-hard. In addition, Ben-Ayed and Blair (1990) declared that a good exact algorithm to solve BLP problems is unlikely to exist by proving that the BLP problem is NP-hard. Moreover, other researchers such as Hansen et al. (1992) and Vicente et al. (1994) also proved that BLP problems are NP-hard and stated that it is difficult to find a globally optimal solution.

Bard (2013) categorized the algorithms to solve bilevel linear programming problems into three different approaches. However, most of these algorithms are not widely applicable due to computational limitations and simplifying assumptions, which they need. The first method uses some form of vertex enumeration, the second one applies a penalty approach to convert the lower-level problem to an unconstrained mathematical program, and the third one involves the Karush-Kuhn-Tucker (KKT) approach to convert the bilevel programming problem to a single level problem [Bard (2013)]. The third category is the most popular method to solve BLP problems, and the most commonly used algorithms in this category are Branch and Bound, Big-M, and Benders algorithms. A branch and bound algorithm to solve BLP problems was proposed by Bard and Moore (1990). They assumed that the feasible region is nonempty and compact,

and they converted the bilevel problem to a single level mathematical problem with complementarity constraints by applying KKT conditions. Then, they used a branch and bound method to deal with the complementarity constraints. Another algorithm to deal with complementarity constraints is the big-M method. Fortuny-Amat and McCarl (1981) reformulated the problem with complementarity constraints as a mixed-integer linear programming problem by adding a binary variable and a large enough positive constant M. The new reformulated mixed-integer problem can be solved by current solvers and algorithms. However, it is hard to know how large a value for M is sufficient. A too small M can eliminate the optimal solution, and a too large M may cause computational errors. Therefore, using this algorithm has practical issues and difficulties. Hu et al. (2008) developed a big-M-free algorithm based on a logical Benders scheme to solve a linear programming model with complementarity constraints. Therefore, their algorithm can also be used to solve BLP problems after applying KKT conditions and converting it to a single level problem with complementarity constraints.

We have developed a new branch and bound algorithm which is more efficient than the current ones. The relaxation problem, which is solved in each node of the algorithm introduced by Bard and Moore (1990), is subdivided into two smaller problems in our algorithm. Therefore, it shortens the solving time of the relaxation problem in each node. The results of 100 randomly generated instances with different sizes show that the new proposed algorithm can solve the bilevel linear programming problems faster than the branch and bound method proposed by Bard and Moore (1990).

The remainder of the dissertation is organized as follows. The first paper on studying the interactions between two major renewable energy policies is presented in Chapter 2 and has been published in *Applied Energy*. In Chapter 3, we present the second paper on developing a tri-level model for inventory control optimization. This paper has been submitted to *European Journal of Operational Research*. In Chapter 4, we introduce a new branch and bound algorithm to solve bilevel linear programming models as the third paper; it is in preparation for submission to *IIE Transactions*. Finally, Chapter 5 concludes the dissertation and proposed possible future research directions.

BIBLIOGRAPHY

Bard, J. F. (1991). Some properties of the bilevel programming problem. *Journal of Optimization Theory and Applications*, 68(2):371–378.

Bard, J. F. (2013). *Practical bilevel optimization: algorithms and applications*, volume 30. Springer Science & Business Media.

Bard, J. F. and Moore, J. T. (1990). A branch and bound algorithm for the bilevel programming problem. *SIAM Journal on Scientific and Statistical Computing*, 11(2):281–292.

Bard, J. F., Plummer, J., and Sourie, J. C. (2000). A bilevel programming approach to determining tax credits for biofuel production. *European Journal of Operational Research*, 120(1):30–46.

Beaudin, M. and Zareipour, H. (2015). Home energy management systems: A review of modelling and complexity. *Renewable and Sustainable Energy Reviews*, 45:318–335.

Ben-Ayed, O. and Blair, C. E. (1990). Computational difficulties of bilevel linear programming. *Operations Research*, 38(3):556–560.

Camacho-Vallejo, J.-F., González-Rodríguez, E., Almaguer, F.-J., and González-Ramírez, R. G. (2015). A bi-level optimization model for aid distribution after the occurrence of a disaster. *Journal of Cleaner Production*, 105:134–145.

Chen, B., Wang, J., Wang, L., He, Y., and Wang, Z. (2014). Robust optimization for transmission expansion planning: minimax cost vs. minimax regret. *IEEE Transactions on Power Systems*, 29(6):3069–3077.

Chiou, S.-W. (2005). Bilevel programming for the continuous transport network design problem. *Transportation Research Part B: Methodological*, 39(4):361–383.

Dempe, S., Kalashnikov, V., and Rıos-Mercado, R. Z. (2005). Discrete bilevel programming: Application to a natural gas cash-out problem. *European Journal of Operational Research*, 166(2):469–488.

Fortuny-Amat, J. and McCarl, B. (1981). A representation and economic interpretation of a two-level programming problem. *The Journal of the Operational Research Society*, pages 783–792.

Hansen, P., Jaumard, B., and Savard, G. (1992). New branch-and-bound rules for linear bilevel programming. *SIAM Journal on Scientific and Statistical Computing*, 13(5):1194–1217.

Hu, J., Mitchell, J. E., Pang, J.-S., Bennett, K. P., and Kunapuli, G. (2008). On the global solution of linear programs with linear complementarity constraints. *SIAM Journal on Optimization*, 19(1):445–471.

Jin, S. and Ryan, S. M. (2014). A tri-level model of centralized transmission and decentralized generation expansion planning for an electricity market, part I. *IEEE Transactions on Power Systems*, 29(1):132–141.

Kuo, R. and Han, Y. (2011). A hybrid of genetic algorithm and particle swarm optimization for solving bi-level linear programming problem–a case study on supply chain model. *Applied Mathematical Modelling*, 35(8):3905–3917.

Lu, J., Shi, C., and Zhang, G. (2006). On bilevel multi-follower decision making: General framework and solutions. *Information Sciences*, 176(11):1607–1627.

Sakawa, M. and Nishizaki, I. (2009). *Cooperative and noncooperative multi-level programming*, volume 48. Springer Science & Business Media.

Saranwong, S. and Likasiri, C. (2016). Product distribution via a bi-level programming approach: Algorithms and a case study in municipal waste system. *Expert Systems with Applications*, 44:78–91.

Tookanlou, M., Ardehali, M., and Nazari, M. (2015). Combined cooling, heating, and power system optimal pricing for electricity and natural gas using particle swarm optimization

based on bi-level programming approach: Case study of canadian energy sector. *Journal of Natural Gas Science and Engineering*, 23:417–430.

Vicente, L., Savard, G., and Júdice, J. (1994). Descent approaches for quadratic bilevel programming. *Journal of Optimization Theory and Applications*, 81(2):379–399.

CHAPTER 2. POTENTIAL COMPETITION FOR BIOMASS BETWEEN BIOPOWER AND BIOFUEL UNDER RPS AND RFS2

A paper published in Applied Energy

Mohammad Rahdar, Lizhi Wang, and Guiping Hu

Abstract

Driven by Renewable Portfolio Standards and Renewable Fuel Standard, biopower generation and biofuel production will increasingly compete for the same biomass resource over the next two decades. We use a linear programming model to study this competition as well as other interactions between the two policies. Our model describes the U.S. renewable energy portfolio by explicitly accounting for all major renewable energy resources, unique resource availability and policy requirements in all 50 states and Washington D.C., and policy deadlines set by all RPS and RFS2 policies within a 2013-2035 modeling horizon. Our modeling results were used to address five important questions regarding interactions between RPS and RFS2 and the impact on U.S. renewable energy portfolio.

2.1 Introduction

Renewable Portfolio Standard (RPS) and the revised Renewable Fuel Standard (RFS2) are expected to be two major policy drivers for the growth of the renewable energy portfolio in the United States in the next couple of decades. Although numerous studies have been conducted to assess these policies separately, most focused on their effectiveness in fostering the growth of a subset of the renewable energy portfolio within a geographic region defined in the policy jurisdiction, and few have examined the interactions between RPS and RFS2 or the implications of such interactions on the nation's holistic renewable energy portfolio. In particular, biomass

can be used to either generate electricity (biopower) to meet the RPS mandates or to produce biofuel to meet the RFS2 requirement. As such, the two policies have created an incentive for biopower and biofuel to compete for the same resource. However, the short-term outcome and long-term implications of such competition have yet to be fully understood by policy makers and other stakeholders of the renewable energy industry. Therefore, we are motivated to examine the potential competition for biomass between biopower and biofuel, other interactions between RPS and RFS2, and the implication of these interactions on the growth of the U.S. renewable energy portfolio over the next two decades.

To understand the status quo of the U.S. renewable energy portfolio, we created a diagram using data from Table A17 of the EIA Annual Energy Outlook 2012 [AEO (2012)], as shown in Figure 2.1. Biomass was the resource for 51% of the total renewable energy consumed in 2012 (8.4 quadrillion Btu)[1] in five sectors: residential 6%, commercial 1%, industrial 26% (collectively referred to as R.C.I.), transportation (biofuel) 15%, and biopower 3%. Wind 15%, geothermal 2%, and solar 0.4% (collectively referred to as W.G.S.) accounted for 17% and hydropower 32% of the total renewable energy portfolio.

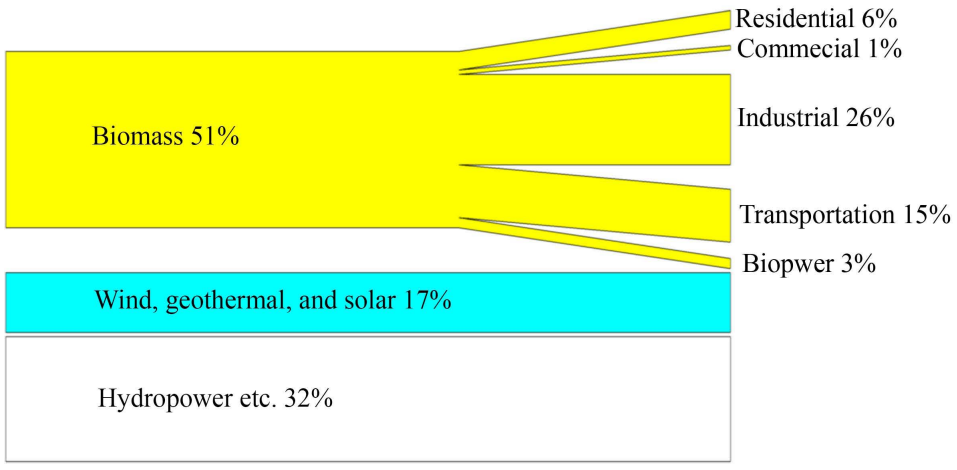


Figure 2.1: The U.S. renewable energy portfolio in 2012. Data from AEO (2012)

The RPS and RFS2 policy drivers, along with others such as the production tax credits (PTC) or investment tax credits (ITC), will drive the U.S. renewable energy portfolio in 2035 very different than it was in 2012. RPS targets on increasing renewable electricity, includ-

[1]One Btu equals to 1055.0559 joules

ing biopower, W.G.S. power, and hydropower [Wiser et al. (2007)]. As of April 2013, thirty states have established RPS mandates and eight have set similar but non-binding goals [DSIRE (2013d)]. So far, the RPS rules in different states are all unique. These rules differ by program structure, enforcement mechanism, classification of generating technologies in tiers, mandated percentages or MWh of renewable electricity generation, deadlines, and non-compliance penalties. There is a rich body of literature on the feasibility and potential impact of RPS.

Johnson and Moyer (2012) analyzed the Illinois RPS and suggested that full implementation of the legislation in Illinois (and perhaps other states) is unlikely without "continued reductions in wind and solar costs and/or an unforeseen rise in wholesale electricity rates". Cory and Swezey (2007) discussed the "hurdle of RPS rules that vary from state to state" that implementation of RPS must surmount to be successful. Carley (2009) found that "states with RPS policies do not have statistically higher rates of RE [renewable energy] share deployment than states without RPS policies." On the contrary, Yin and Powers (2010) used a new measure of policy stringency to argue that "RPS policies have had a significant and positive effect on in-state renewable energy development". They also pointed out that allowing for free trade of renewable energy certificates "can significantly weaken the impact of an RPS". Menza and Vachon (2006) also found RPS to be effective in "promoting the development of wind capacity." Palmer and Burtraw (2005) compared the cost effectiveness of RPS, production tax credit, and cap-and-trade and concluded that cap-and-trade is more effective in achieving carbon emission reductions than the other two.

Renewable Fuel Standard (RFS) is a federal program designed to help protect public health and the environment and reduce the dependence on imported petroleum. Renewable fuels are defined as liquid or gaseous fuels derived from renewable biomass energy sources. A mandatory minimum volume of biofuel to be used in the national transportation fuel supply was established in 2005 with the Energy Policy Act. The initial standard mandated that the minimum usage volume of renewable fuel rise to 7.5 billion gallons[2] by 2012. Two years later, the Energy Independence and Security Act of 2007 expanded the biofuel mandate to 36 billion gallons of (including 16 billion gallons for cellulosic and 20 billion gallons for non-cellulosic) biofuel to be

[2]One U.S. liquid gallon equals to 0.0038 cubic meter

blended into transportation fuel by 2022 [Schnepf and Yacobucci (2010)]. This revised RFS is referred to as RFS2. A few recent studies have started to address the potential interactions of RFS2 with other policies. Jeffers et al. (2013) studied the bioenergy feedstock commodity market with three buyers: biopower, biofuel, and foreign exports. Their simulation model showed that either biofuel or overseas biomass demand could dominate the market under different policy settings. They also suggested that market competition can "effectively drive up prices for the biomass feedstocks and potentially exclude industries from the market". Huang et al. (2013) studied the interactions of three policies: RFS2, low carbon fuel standard (LCFS), and a carbon price. They concluded that "the addition of a LCFS to the RFS increases the share of second generation biofuel; the addition of a carbon price to these policies encourages fuel conservation; these combined policies significantly increase the reduction in GHG emissions; [and] they also achieve greater energy security and economic benefits than the RFS alone".

Our study makes a new contribution to the existing literature by pioneering the analysis on the interactions between RPS and RFS2. In particular, we are motivated to seek answers to the following questions that have not been elucidated by previous studies. These questions are difficult to address without looking at how the two policies (along with others) jointly affect the entire renewable energy market along all resource, geographical, and temporal dimensions.

Q1: What are the potential interactions between RPS and RFS2?

Q2: Under RPS and RFS2, how will the competition for biomass between biopower generation and biofuel production progress in the next two decades?

Q3: Under RPS and RFS2, what is the outlook of renewable energy portfolio in the U.S.?

Q4: How will different states' unique renewable energy portfolios evolve in the next two decades?

Q5: What factors is the U.S. renewable energy portfolio most sensitive to?

2.2 Model

In order to address the five questions that motivated this study, we constructed an optimization model to describe the overarching interactions within the complex renewable energy portfolio from resource, geographical, and temporal dimensions. First, we include all major renewable energy resources (biomass, W.G.S., and hydro) and demand sectors (biopower, non-cellulosic and cellulosic biofuel, W.G.S. power, and hydropower) into the modeling framework. As such, the prediction of renewable energy portfolio from our model resulted from careful evaluation of costs (capital investment cost, operating and maintenance costs, and non-compliance penalties) and benefits (sales revenue and tax credits) of each technology rather than over-simplifying presumptions. Second, our model treats all 50 states and Washington D.C. as 51 separate entities, each having their own reserves of renewable energy resources and unique RPS requirements (mandates, goals, or neither). Nevertheless, our model also captures the interactions among different states, including truck transportation of biomass and RFS2 compliances. Third, we use a 23-year modeling horizon, which allows us to accommodate practical considerations of market trends before and after RPS and RFS2 deadlines as well as time value of money.

We made several major simplifying assumptions, some of which are due to lack of good data and others are believed to be necessary to maintain tractability of the model without significantly compromising the credibility of the results. First, our optimization model adopts a centralized and coordinative planning perspective by maximizing the net present value of the total profit (benefits less costs) of the U.S. renewable energy industry, which is used to approximate the investment and operating decisions for all states across all renewable energy sectors throughout the modeling horizon. In reality, investment and operating decisions are made by multiple decision makers in electricity and transportation fuel markets to serve their own objectives, some competitively and others in coordination. Thus, game theoretic models would be able to better describe such market behavior. However, game theoretic models would not only require much higher modeling granularity and more sophisticated database but also encounter much more complicated computational challenges such as the tractability, existence,

and uniqueness of a market equilibrium. Our optimization model avoids such problems by assuming that the invisible hand of economy will direct the overall flow of capital and natural resources in the most efficient manner towards cost minimization and profit maximization for the entire industry. Second, our model is deterministic, not taking uncertainty into explicit consideration. To address the concerns raised in Q5 regarding uncertainty and its potential impact on the renewable energy output, we conduct a sensitivity analysis by examining the impact of dozens of parameters on the results. Third, our model treats several factors as known parameters rather than decision variables due to their lack of unforeseeable interactions with the rest of the model. For example, demands of biomass energy in the R.C.I. sectors are not directly affected by either RPS or RFS2, thus their projections in the next two decades are treated as known. Non-cellulosic biofuel (mostly corn ethanol and soybean diesel) production is also assumed to exactly meet the RFS2 requirement due to abundant existing capacity of these conventional biofuel production facilities. Fourth, we do not treat hydropower as RPS eligible for any state. Since the goal of the RPS is to encourage new investment in renewable energy, and most hydroelectric facilities were installed decades ago, most states place certain restrictions on hydropower by capacity, vintage, or technology, and some do not count hydropower at all. Some legislations regarding the RPS eligibility of hydropower are difficult to formulate in the model or require more detailed data than what is publicly available.

Along the resource dimension of our optimization model as described above, the structure of the model is depicted in Figure 2.2, which exactly represents the major resources and demand sectors of the U.S. renewable energy portfolio as diagramed in Figure 2.1. Following the categorization in Haq and Easterly (2006), we consider four major types of biomass: agricultural residues, energy crops, forestry residues, and urban wood waste. The "other" category mostly accounts for conventional biomass resources such as corn or soybean. Due to the aforementioned reasons, the R.C.I. sectors, non-cellulosic biofuel, and hydropower are treated as known parameters (all colored in blue) and are not formulated as decision variables in the model. To accurately incorporate RPS policy, our model sets a separate constraint for each RPS state and for each eligible renewable energy defined in the legislation. Non-compliance penalties for

different types of renewable energy in different states are also captured in the model. The RFS2 policy is similarly formulated as a soft constraint with a penalty for non-compliances.

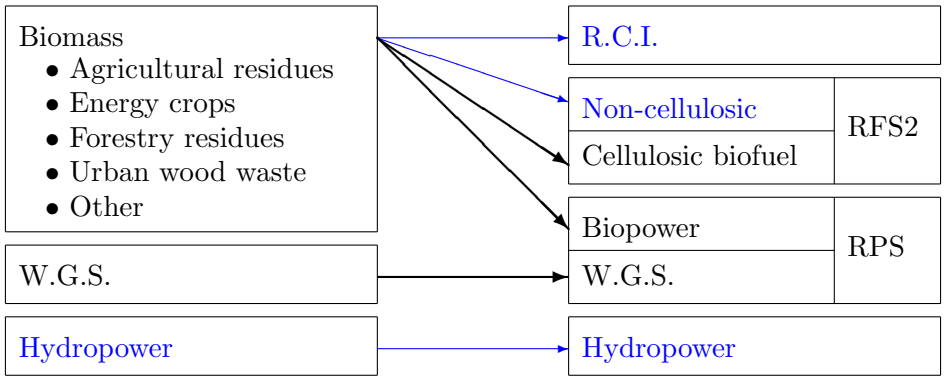


Figure 2.2: Modeling structure

Using the sets, parameters, and decision variables defined in Appendix A, the mathematical formulation of our optimization model is presented as follows.

$$\max \quad \zeta = \sum_{v,j,t}(1+r)^{(t_0-t)}(\beta_{j,t}+\varphi_{v,t})x_{v,j,t} + \sum_{j,t}(1+r)^{(t_0-t)}(\beta^{\mathrm{F}}_t+\varphi^{\mathrm{F}}_t)x^{\mathrm{F}}_{j,t} \tag{2.1}$$

$$-\sum_{u,j,t}(1+r)^{(t_0-t)}c_{u,j,t}x_{u,j,t} - \sum_{u,i\neq j,t}(1+r)^{(t_0-t)}\pi_{i,j,t}y_{u,i,j,t} \tag{2.2}$$

$$-\sum_{v,j,t}(1+r)^{(t_0-t)}(c_{v,j,t}+f_{v,j,t})x_{v,j,t} - \sum_{j,t}(1+r)^{(t_0-t)}c^{\mathrm{F}}_{j,t}x^{\mathrm{F}}_{j,t} \tag{2.3}$$

$$-\sum_{v,j,t}(1+r)^{(t_0-t)}l_{v,j,t}(1-\lambda_{v,t})z_{v,j,t} - \sum_{j,t}(1+r)^{(t_0-t)}l^{\mathrm{F}}_{j,t}(1-\lambda^{\mathrm{F}}_t)z^{\mathrm{F}}_{j,t} \tag{2.4}$$

$$-\sum_{j,t,k}(1+r)^{(t_0-t)}\mu_{j,t,k}s_{j,t,k} - \sum_{t}(1+r)^{(t_0-t)}\mu^{\mathrm{F}}_t s^{\mathrm{F}}_t \tag{2.5}$$

$$\text{s. t.} \qquad \sum_u \rho_u \left(x_{u,j,t} + \sum_{i\neq j} y_{u,i,j,t} - \sum_{i\neq j} y_{u,j,i,t}\right)$$

$$\geq d_{j,t} + 1.45\times 10^{-2}x_{\mathrm{biomass},j,t} + 2.90\times 10^{-4}x^{\mathrm{F}}_{j,t} \qquad \forall j,t \tag{2.6}$$

$$x_{u,j,t} + \sum_{i\neq j} y_{u,i,j,t} \geq \sum_{i\neq j} y_{u,j,i,t} \qquad \forall u,j,t \tag{2.7}$$

$$p_{v,j,t} = p_{v,j,(t-1)} + z_{v,j,t} \qquad \forall v,j,t \tag{2.8}$$

$$p^{\mathrm{F}}_{j,t} = p^{\mathrm{F}}_{j,(t-1)} + z^{\mathrm{F}}_{j,t} \qquad \forall j,t \tag{2.9}$$

$$x_{u,j,t} \leq p_{u,j,t} \qquad \forall u,j,t \tag{2.10}$$

$$x_{v,j,t} \leq 8760\alpha_v p_{v,j,t} \qquad \forall v,j,t \tag{2.11}$$

$$x^{\mathrm{F}}_{j,t} \leq p^{\mathrm{F}}_{j,t} \qquad \forall j,t \tag{2.12}$$

$$z_{v,j,t} \leq M_{v,j,t} \qquad \forall v, j, t \tag{2.13}$$

$$z^{\mathrm{F}}_{j,t} \leq M^{\mathrm{F}}_{j,t} \qquad \forall j, t \tag{2.14}$$

$$\textstyle\sum_v q_{v,j,k} x_{v,j,t} + s_{j,t,k} \geq \eta_{j,t,k} e_{j,t} \qquad \forall j, t, k \tag{2.15}$$

$$\textstyle\sum_j x^{\mathrm{F}}_{j,t} + s^{\mathrm{F}}_t \geq \theta_t \qquad \forall t \tag{2.16}$$

$$\text{all decision variables} \geq 0 \tag{2.17}$$

The objective function (2.1)-(2.5) of the model is to maximize the net present value of the total profit (revenue less cost) of the renewable energy industry. In (2.1), the first term is the total revenues from sales (β) and production tax credits (φ) for W.G.S. power and biopower generation (x), and the second term is revenue for cellulosic biofuel production. The discount factor r is used to calculate the present value of future cash flows. The eight cost terms in (2.2)-(2.5) are for, respectively, biomass production, biomass transportation, renewable electricity generation (variable cost c plus fixed cost f), biofuel production, capital investment (adjusted by investment tax credit) in new renewable power plants, capital investment (adjusted by investment tax credit) in biofuel production facilities, penalties for RPS non-compliances, and penalties for RFS2 shortfalls. Constraint (2.6) requires that the amount of biomass production and imports minus exports must exceed demand from R.C.I. sectors, biopower generation, and biofuel production (all converted to BBtu). Constraint (2.7) sets the combined amount of biomass production and imports as the upper limit for exports. Equations (2.8) and (2.9) update the yearly capacities of renewable electricity generation (in MW) and biofuel production (in gallon) to account for new additions. Constraints (2.10)-(2.12) define the available capacity for biomass production, renewable electricity generation, and biofuel production, respectively. Constraints (2.13) and (2.14) set the upper bounds of new capacities for investment in renewable power plants and biofuel production facilities that can be realistically put in due to limitations in manufacturing capability, resource (material, labor, funds, etc.) availability, and legislative requirements. Constraints (2.15) and (2.16) set RPS and RFS2 requirements. The RFS2 target is an aggregate for all states, whereas RPS mandates are specified for each state and each type of renewable energy. The binary parameter $q_{v,j,k}$ indicates whether or not renewable electricity

type v is included in tier k of state j's RPS legislation. All decision variables are required to be non-negative in Constraint (2.17).

2.3 Results

The linear program model (2.1)-(2.17) contains 261,603 decision variables and 22,380 constraints. The entire data set take more than 1 MB of hard drive space. It was programmed in GAMS and solved to optimality in a few seconds on a desktop computer with standard configurations. Data used for all sets and parameters in the model are explained in Appendix B. We present our modeling results by answering the five motivating questions.

Q1: What are the potential interactions between RPS and RFS2?

A1: We assess the potential interactions between RPS and RFS2 by comparing the modeling results with four cases of policy implementation: no policy (case 1), RPS only (case 2), RFS2 only (case 3), and both policies (case 4). Numerical results are summarized in Table 2.1. Without RFS2, RPS would increase 65.19 billion kWh/year of W.G.S. power and 25.85 billion kWh/year of biopower, averaged between 2013 and 2035. This effect represents an increase of the nationwide renewable electricity portfolio (excluding hydropower) from 6.87% in 2013 to 11.47% in 2035. On the other hand, without RPS, RFS2 would increase nationwide cellulosic biofuel production by an average of 7.69 billion gallons/year. The interaction of the two policies reduces the contributions of both. Specifically, due to the competition for biomass from RFS2, a yearly average of 7.70 million tons of biomass that would have been used to generate biopower under RPS will be used to produce cellulosic biofuel instead. Reversely, due to the competition for biomass from RPS, a yearly average of 5.01 million tons of biomass that would have been used to produce cellulosic biofuel under RFS2 will be used to generate biopower instead. We also point out that the interactions between RPS and RFS2 have little impact on W.G.S.; they only affect the total amount of biomass production and the allocation of the biomass resource for biopower and cellulosic biofuel.

Table 2.1: Modeling results in four cases

	Case 1	Case 2	Case 3	Case 4
RPS RFS2		✓	 ✓	✓ ✓
Average W.G.S. power generation (billion kWh/year) from 2013 to 2035	320.68	385.87	320.68	386.01
Average biopower generation (billion kWh/year) from 2013 to 2035	0.17	26.01	0.17	16.78
Average biomass used for biopower generation (million ton/year)	0.14	21.68	0.14	13.98
Average cellulosic biofuel production (billion gallon/year) from 2013 to 2035	0.02	0.02	7.71	7.41
Average biomass used for cellulosic biofuel production (million ton/year)	0.32	0.32	128.54	123.53

Q2: Under RPS and RFS2, how will the competition for biomass between biopower generation and biofuel production progress in the next two decades?

A2: To address this question, we plot in Figure 2.3 the projection of four sectors of renewable energy consumption in the U.S. that are based on biomass resources. The R.C.I. projection is adopted from AEO (2012), the non-cellulosic biofuel production is assumed to exactly meet the RFS2 requirements, and the projections for biopower and cellulosic biofuel are from our modeling results. The figure shows that biomass based renewable energy will increase by 69% in the next two decades, with R.C.I. and non-cellulosic biofuel accounting for a combined 95% and 87% in 2013 and 2035, respectively. The competition for biomass between biopower and biofuel is expected to turn sharply from biopower being the dominating pathway to the opposite. Biopower is expected to shrink by 67% over the next two decades due to lack of cost competitiveness compared to W.G.S. power generation technologies as well as the distraction from RFS2. This result is consistent with the findings from Dassanayake and Kumar (2012) in which triticale straw-based biopower generation is less economically competitive than coal-based electricity generation. Driven by RFS2, annual production of cellulosic biofuel is expected to surge from 0.14 billion gallons in 2013 to 8.91 billion gallons in 2022 (7.09 billion gallons short of the

16 billion-gallon target) and then 8.81 billion gallons in 2035. The downturn of cellulosic biofuel production after 2023 is due to the assumed expiration of the cellulosic biofuel producer tax credit in 2022.

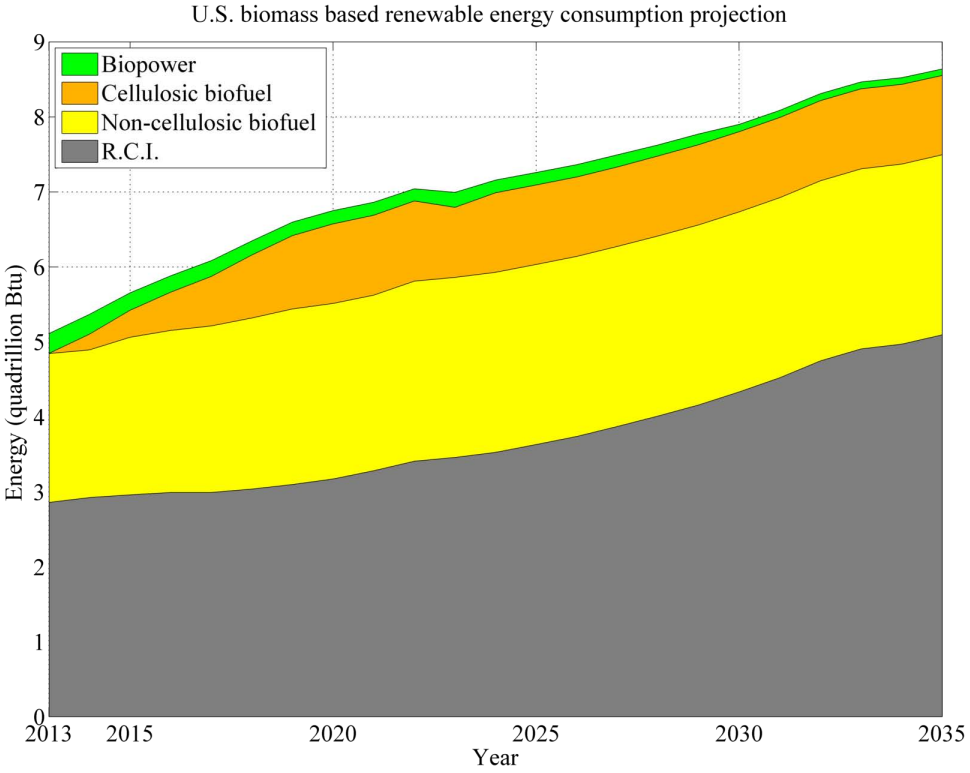


Figure 2.3: Projection of biomass based renewable energy consumption in the U.S. Results from our model are combined with data from AEO (2012).

Q3: Under RPS and RFS2, what is the outlook of renewable energy portfolio in the U.S.?

A3: Table A17 in the Annual Energy Outlook 2012 [AEO (2012)] as well as our modeling results can be used to address this question. In Figure 2.4, we plot the EIA projection of renewable energy consumption broken into seven categories. For the purpose of model validation, we also plot the same seven categories of projection with an additional differentiation of cellulosic and non-cellulosic biofuel from our modeling results in Figure 2.5. Since our model does not include projections for non-cellulosic biofuel, hydropower, and R.C.I., we use the same data for those sectors from AEO (2012) in Figure 2.5. The overall trend of our projections is consistent with the EIA results. However, we are more

optimistic than EIA on the growth of W.G.S. power but less so on biopower. In fact, EIA expects biopower to grow 2.4-fold between 2013 and 2035, whereas we predict a 67% shrink. Moreover, we are more optimistic than EIA about the growth of cellulosic biofuel production before the 2022 deadline, but we expect the production to stay at the same level with a slight fallback afterwards rather than continuing to grow throughout 2035 as EIA projected. According to Annual Energy Outlook 2012, 22.1 billion gallons of biofuel (including cellulosic and non-cellulosic) will be produced in 2022, which is 13.9 billion gallons short of the RFS2 target. We predicted 28.91 billion gallons biofuel production in 2022, including 8.91 billion gallons cellulosic (by modeling results) and 20 billion gallons non-cellulosic (by assumption) biofuels, which is 7.09 billion gallons short of the target.

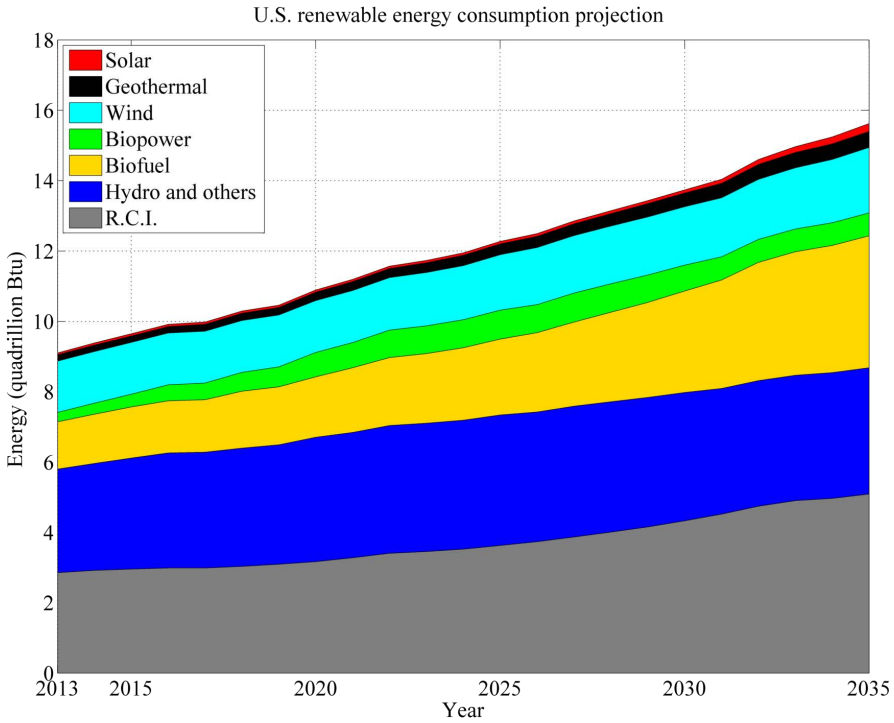


Figure 2.4: Projection of renewable energy consumption in the U.S. from Table A17 in AEO (2012) for the reference case

Q4: How will different states' unique renewable energy portfolios evolve in the next two decades?

A4: Figures 2.6-2.10 show the trends of top states in wind, geothermal, solar, biopower, and cellulosic biofuel, respectively. Each curve represents the trajectory of a certain type of

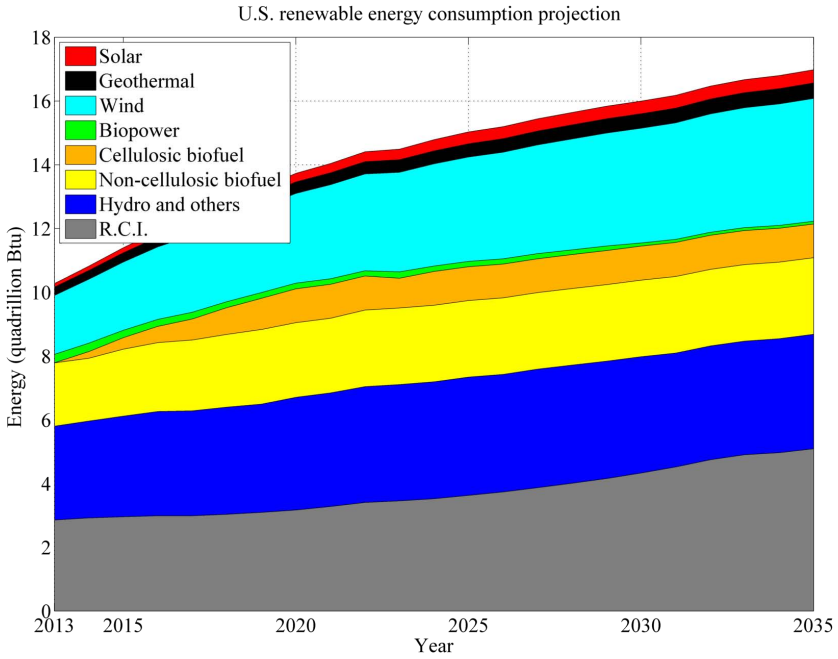


Figure 2.5: Projection of renewable energy consumption in the U.S. from our model under the base case scenario combined with partial data from AEO (2012).

renewable energy generation between 2013 and 2035 in a specified state. Whereas most states show an increasing trend of renewable energy generation, biopower is shrinking and losing the competition to cellulosic biofuel. Figure 2.11 plots the renewable energy portfolios of top 30 renewable energy generating states, which is broken into four types of resources: wind, geothermal, solar, and biomass (for biopower and cellulosic biofuel). Non-cellulosic biofuel or hydropower is not included in Figure 2.11.

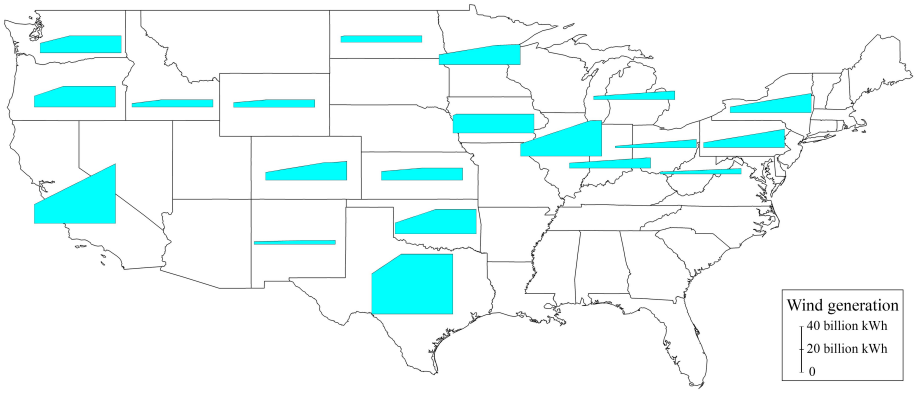


Figure 2.6: Wind power generation in top 20 states between 2013 and 2035.

Figure 2.7: Geothermal power generation in top 10 states between 2013 and 2035.

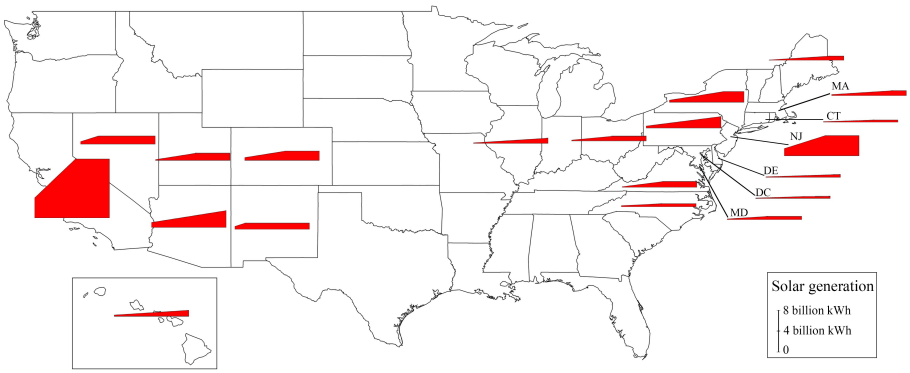


Figure 2.8: Solar power generation in top 20 states between 2013 and 2035.

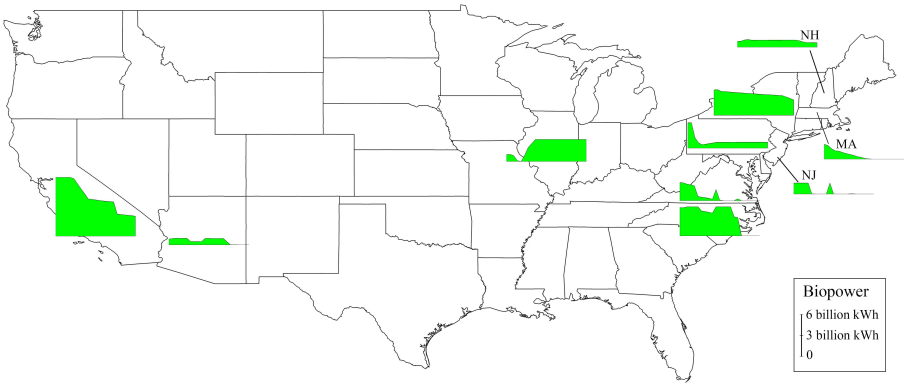


Figure 2.9: Biopower generation in top 10 states between 2013 and 2035.

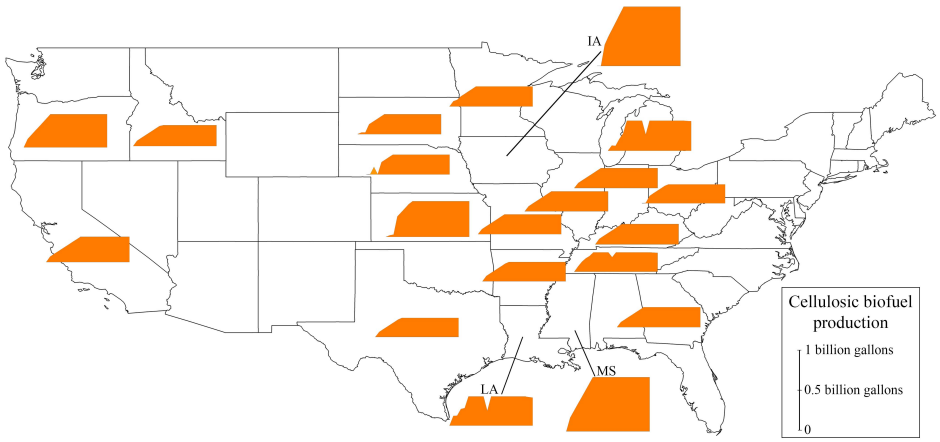


Figure 2.10: Cellulosic biofuel production in top 20 states between 2013 and 2035.

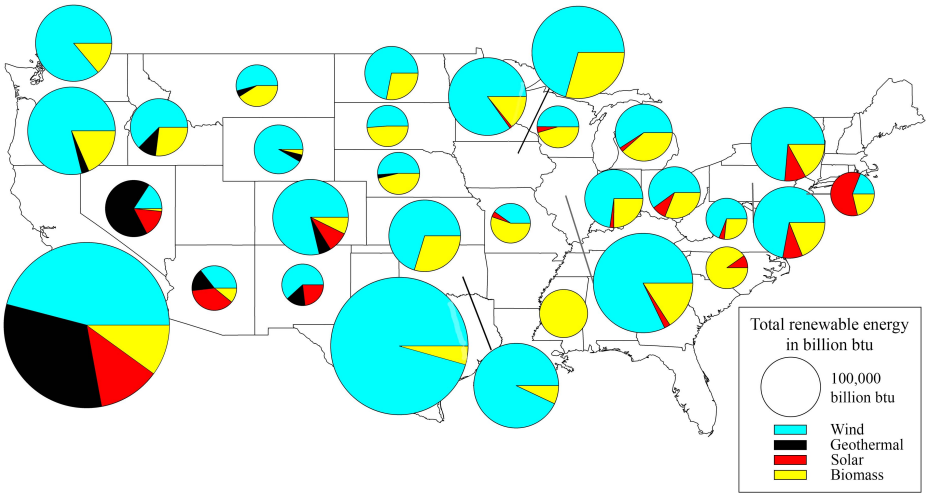


Figure 2.11: Renewable energy portfolios in top 30 states broken into four types of resources averaged between 2013 and 2035, excluding biomass for non-cellulosic biofuel production.

Q5: What factors is the U.S. renewable energy portfolio most sensitive to?

A5: To quantify the sensitivity of renewable energy production with respect to multiple parameters, we define three scenarios each comprising of a set of values for these parameters: base case, optimistic, and pessimistic scenarios. The optimistic and pessimistic scenarios are defined in such a way that the nationwide total renewable energy generation would be increased and decreased with respect to the base case, respectively. The objective of

this analysis is to identify parameters that would have the most significant impact on the modeling results. Answers A1-A4 were all based on the base case scenario, which we believe represents the most likely realization of the uncertain parameters. Parameter values for the base case scenario are described in Appendix B. The changes of parameter values for these two scenarios are described as follows.

Optimistic scenario: Seven cost parameters ($c_{u,j,t}, c_{v,j,t}, c^{\mathrm{F}}_{j,t}, \pi_{i,j,t}, f_{v,j,t}, l_{v,j,t}, l^{\mathrm{F}}_{j,t}$) are reduced by 20% with respect to the base case. Two revenue parameters ($\beta_{j,t}, \beta^{\mathrm{F}}_t$) are increased by 20%. The investment limits $M_{v,j,t}$ and $M^{\mathrm{F}}_{j,t}$ are increased by 20%. Two penalty parameters for RPS and RFS2 non-compliances ($\mu_{j,t,k}, \mu^{\mathrm{F}}_t$) are increased by 50%. Four expiration dates of investment and production tax credits ($\varphi_{v,t}, \varphi^{\mathrm{F}}_t, \lambda_{v,t}, \lambda^{\mathrm{F}}_t$) are all extended to 2035.

Pessimistic scenario: The seven cost parameters are increased by 20% and the two revenue parameters are reduced by 20%. The investment limits $M_{v,j,t}$ and $M^{\mathrm{F}}_{j,t}$ are decreased by 20%. The two penalty parameters are reduced by 50%. The four expiration dates of tax credits ($\varphi_{v,t}, \varphi^{\mathrm{F}}_t, \lambda_{v,t}, \lambda^{\mathrm{F}}_t$) are expedited to 2023, 2013, 2016, and 2013, respectively. These are the expiration dates set by the current regulations, assuming no further extensions.

Our sensitivity analysis results are plotted in Figures 2.12-2.17. In each figure, the effects of all uncertain parameters on one output in the optimistic and pessimistic scenarios are shown against the base case, and the effects are ranked from the largest on top to the smallest at the bottom. Each bar in Figures 2.12-2.17 is obtained by running the model (2.1)-(2.17) with only one change in the parameter (or set of parameters) specified on the left-hand-side of the figures.

Figure 2.12 shows the sensitivity of the total renewable electricity generation with respect to the uncertain parameters we identified. Four factors could affect the total generation by more than 5%: electricity price, wind generation cost (including investment cost and variable cost), PTC for wind, and investment limit on wind. These results suggest that wind energy will play an important role in shaping the renewable electricity development.

Its economic or technological improvement and policy changes will have more impact than any other type of renewable energy on total renewable electricity generation.

Figure 2.13 suggests that wind power generation is most sensitive to five factors: wind generation cost, electricity price, renewal/expiration of PTC for wind, investment limit of wind power, and RPS penalty. Interestingly, either increasing or decreasing RPS penalty will reduce wind power generation. If RPS penalty is decreased, then all renewable electricity generation will fall. On the other hand, if RPS penalty is increased, then solar power and biopower will increase, as can be seen in Figures 2.15 and 2.16, but wind power will fall. These results demonstrate the interactions between multiple renewable energy resources and technologies in response to policy changes.

Figure 2.14 suggests that geothermal power generation is most sensitive to five factors: geothermal generation cost, investment limit of geothermal power, electricity price, and renewal/expiration of PTC and ITC for geothermal. RPS penalties play a similar role as in wind generation. Favorable changes in wind generation cost and investment limit of wind power also affect geothermal generation, but in the opposite direction as they have on wind generation. This is due to the substitutability of W.G.S. resources in fulfilling RPS requirements.

Figure 2.15 suggests that solar power generation is most sensitive to five factors: renewal/expiration of ITC for solar, solar generation cost, investment limit of solar power, electricity price, and RPS penalty. For solar, the increase (or decrease) of RPS penalty does intuitively increase (or decrease) solar power generation. Favorable changes for competing technologies also have negative effects on solar power.

Figure 2.16 suggests that many factors could significantly affect biopower generation. We make two interesting observations. First, nine factors could increase biopower generation by 80% or more, and seven of them could also decrease the generation by 50% or more. Second, biopower generation is very susceptive to the competition from cellulosic biofuel. All favorable (or non-favorable) changes for cellulosic biofuel will negatively (or positively) affect biopower generation, and four of these factors could decrease (or increase) biopower

generation by above 50% (or 100%). In contrast, the same set of factors have much less effect (below 20%) in the more mature W.G.S. technologies.

Figure 2.17 suggests that biofuel production is most sensitive to five factors: renewal/expiration of PTC for biofuel, biofuel price, biofuel cost, RFS2 penalty, and biomass cost, all of which could increase or decrease cellulosic biofuel production by at least 32% and up to 89%. In contrast to biopower, cellulosic biofuel production is much less sensitive to competition from biopower and other types of renewable energy policies.

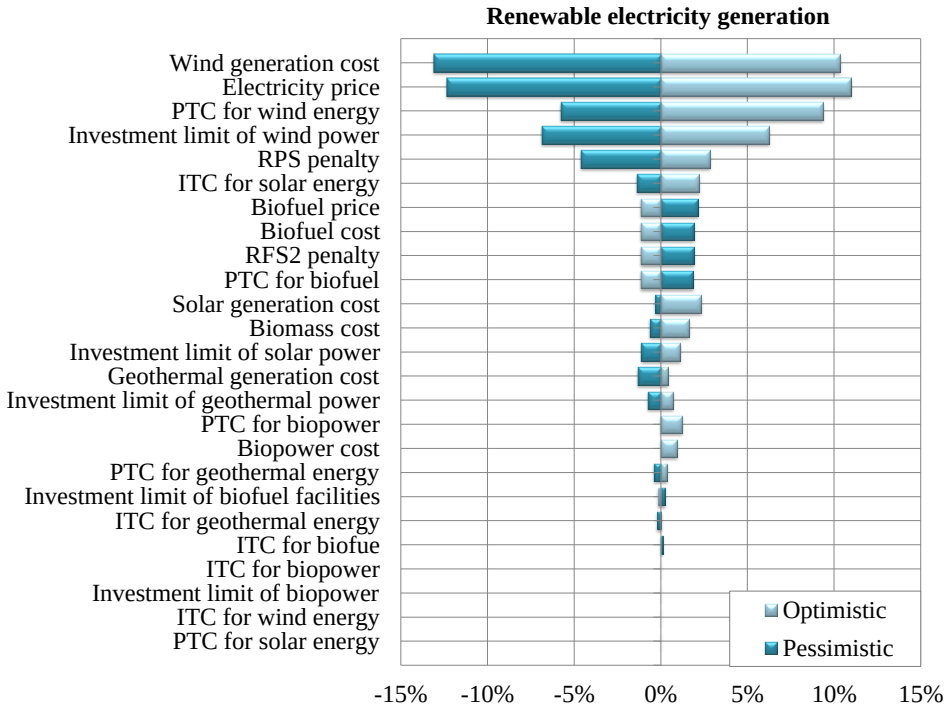


Figure 2.12: Sensitivity of total renewable electricity generation.

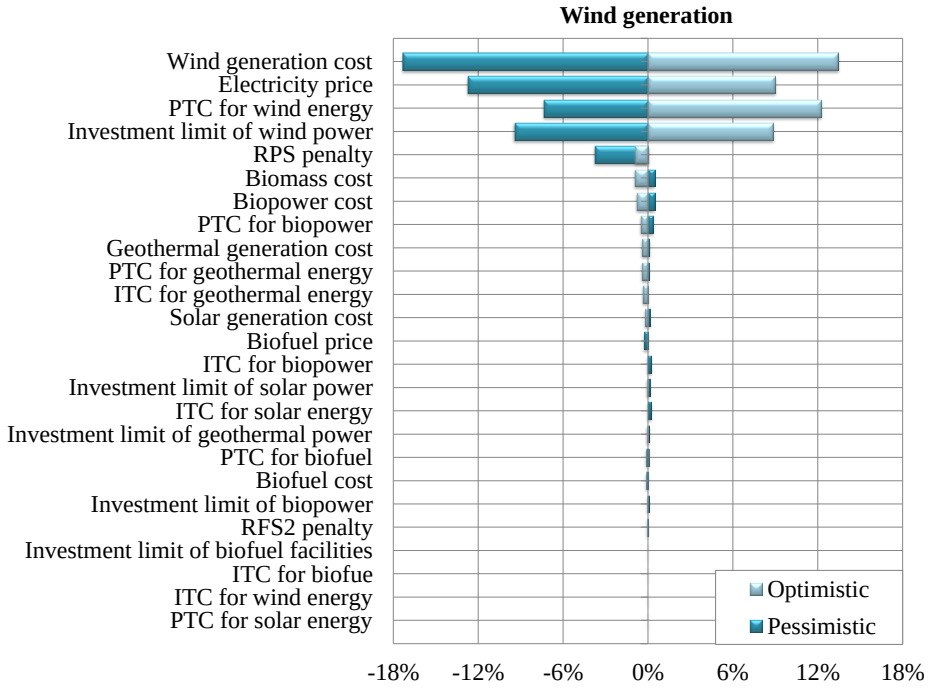


Figure 2.13: Sensitivity of wind power generation.

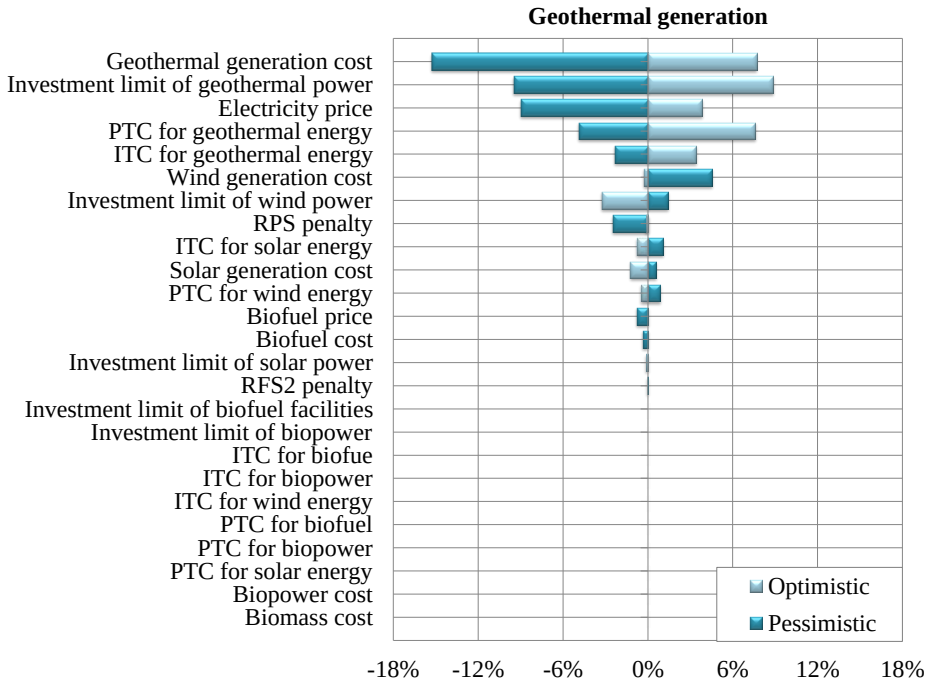


Figure 2.14: Sensitivity of geothermal power generation.

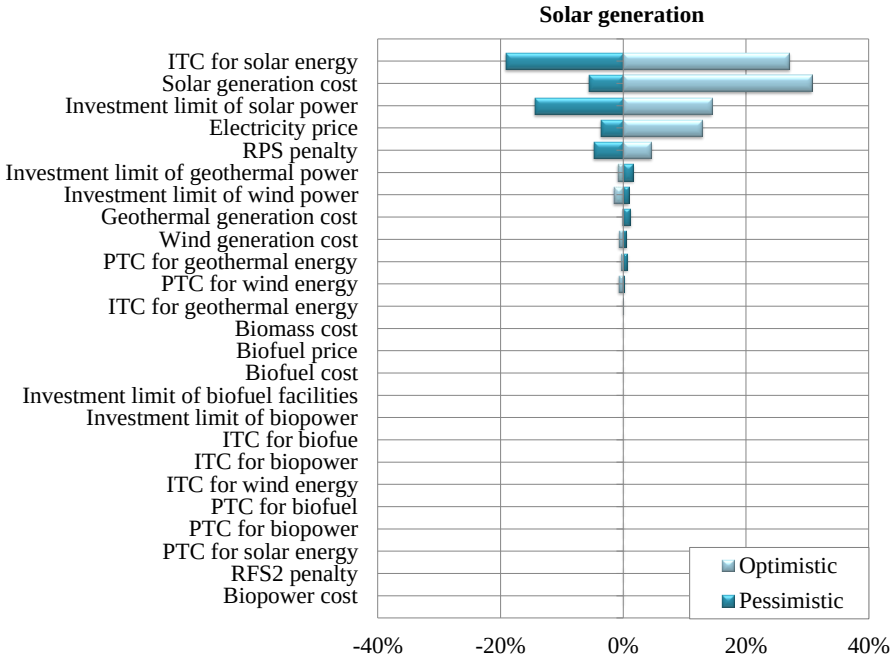


Figure 2.15: Sensitivity of solar power generation.

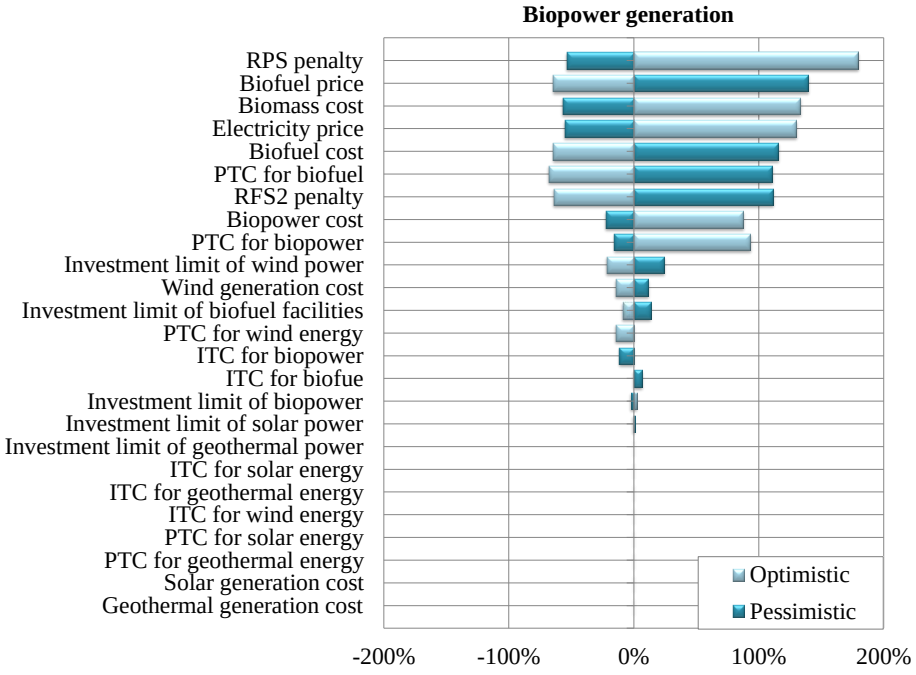


Figure 2.16: Sensitivity of biopower generation.

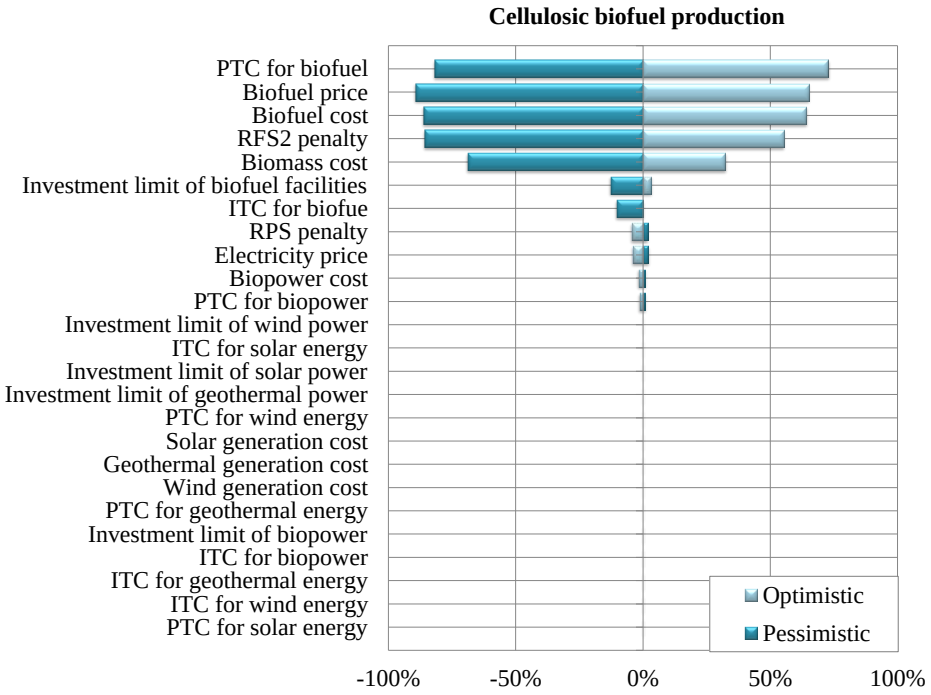


Figure 2.17: Sensitivity of cellulosic biofuel production.

2.4 Conclusions

Our study focused on the potential competition for biomass from RPS driven biopower generation and RFS2 driven biofuel production as well as other interactions between these two policies. As perhaps the first study on this topic, our model has several unique strengths that make it particularly appropriate to address the five important questions Q1-Q5. First, our model takes a systems perspective of the entire renewable energy portfolio. On the resource dimension, availability of multiple renewable energy resources, projection of all major demand sectors in the industry, and investment and operating costs of different generation/production technologies are incorporated. On the geographical dimension, the differences of 50 states and Washington D.C. in renewable energy resource abundance, demand, RPS policies (including different definitions of tiers, deadlines, and penalties), and investment constraints were all explicitly taken into account. On the temporal dimension, a 23-year modeling horizon was used to observe how the U.S. renewable energy industry evolves to pass one deadline after another

set by various RPS and RFS2 legislations. Second, our model is computationally tractable. Efficient linear programming algorithms and software can solve the model to optimality within a few seconds, which allows the model to be solved multiple times to answer what-if questions and for sensitivity analysis. Third, most of the parameters used in our computational study are from publicly available database; when certain data are unavailable, assumptions were carefully made and validated through multiple channels to fill in the gap. Fourth, our computational experiment is conveniently repeatable and extensible for further analysis. All parameters, variables, objective, and constraints of the model are explained; all of the data used as well as their sources are described in Appendices A and B. As a result, improvement can be easily made if additional features of the policy become the focal point of a new research question or more detailed data become available.

Results from our model suggest that cellulosic biofuel production will quickly dominate the competition for biomass against biopower generation. This is because the biomass production and biopower generation costs are higher than those for W.G.S. power, whereas cellulosic biofuel production faces a stringent RFS2 mandate with no cheaper substitution. The renewable energy portfolios in 50 states and Washington D.C. could vary significantly, and they all have their unique trajectories throughout 2035. Our sensitivity analysis reveals that W.G.S. power generation is relatively robust with respect to various uncertain factors, whereas biopower and biofuel are much more susceptible to uncertainty associated with (investment, generation, production) costs, (electricity and transportation fuel) prices, and policies. These analysis results also suggest that the interactions between RPS and RFS2 will have more impact on biopower than on biofuel.

As pointed out in the Introduction section, we made several simplifying assumptions in the model, which may affect the accuracy of our results to some extent. It would be difficult to integrate the strategic behavior of investors in the renewable energy markets without switching to a completely different modeling approach, which may have limitations of its own. However, we expect that more credible results could be obtained by feeding the model with more accurate data, such as investment and operating costs in different states. Moreover, the model can

be extended to incorporate additional features, such as explicit modeling of the eligibility of hydropower in different RPS legislations, given clarification of policy and availability of data.

2.5 Appendix A

Sets

Notation	Definition
$\mathcal{J}$	Set of 50 states in the U.S. and Washington D.C.
$\mathcal{J}^{\text{RPS}}$	Set of 38 U.S. states with RPS (30) or RPG[3](8)
$\mathcal{K}_j$	Set of tiers of RPS policy for state $j \in \mathcal{J}^{\text{RPS}}$
$\mathcal{T}$	Set of years within modeling horizon, $\mathcal{T} = \{t_1, t_2, ..., t_T\}$, where T is the number of years in the modeling horizon.
$\mathcal{U}$	Set of four major types of biomass
$\mathcal{V}$	Set of four major types of renewable electricity resources

Parameters

Notation	Definition	Unit
$c_{u,j,t}$	Biomass production cost of type $u \in \mathcal{U}$ in state $j \in \mathcal{J}$ in year $t \in \mathcal{T}$	\$/ton
$\pi_{i,j,t}$	Biomass transportation cost from state $i \in \mathcal{J}$ to $j \in \mathcal{J}$ in year $t \in \mathcal{T}$	\$/(ton mile)[4]
$c_{v,j,t}/f_{v,j,t}$	Renewable energy generation variable/fix cost of type $v \in \mathcal{V}$ in state $j \in \mathcal{J}$ in year $t \in \mathcal{T}$	\$/MWh
$c^{\text{F}}_{j,t}$	Cellulosic biofuel production cost in state $j \in \mathcal{J}$ in year $t \in \mathcal{T}$	\$/gallon
$l_{v,j,t}$	Capital investment cost of renewable electricity generation of type $v \in \mathcal{V}$ in state $j \in \mathcal{J}$ in year $t \in \mathcal{T}$	\$/MW
$l^{\text{F}}_{j,t}$	Capital investment cost of cellulosic biofuel facilities in state $j \in \mathcal{J}$ in year $t \in \mathcal{T}$	\$/gallon
$\beta_{j,t}$	Average wholesale electricity price in state $j \in \mathcal{J}$ in year $t \in \mathcal{T}$	\$/MWh
β^{F}_t	Average biofuel price in year $t \in \mathcal{T}$	\$/gallon
$\varphi_{v,t}$	Production tax credit for renewable electricity generation of type $v \in \mathcal{V}$ in year $t \in \mathcal{T}$	\$/MWh
φ^{F}_t	Cellulosic biofuel producer tax credit in year $t \in \mathcal{T}$	\$/gallon

[3] Renewable Portfolio Goal
[4] One mile equals to 1.6093 kilometers

$\lambda_{v,t}$	Investment tax credit as a percentage discount of capital investment of renewable electricity generation of type $v \in \mathcal{V}$ in year $t \in \mathcal{T}$	unitless
λ_t^{F}	Investment tax credit as a percentage discount of capital investment of cellulosic biofuel production facilities in year $t \in \mathcal{T}$	unitless
$\mu_{j,t,k}$	Penalty for non-compliance with RPS tier $k \in \mathcal{K}_j$ in state $j \in \mathcal{J}$ in year $t \in \mathcal{T}$	\$/MWh
μ_t^{F}	Penalty for non-compliance with RFS2 in year $t \in \mathcal{T}$	\$/gallon
r	Discount rate	unitless
ρ_u	Conversion factor of 1 ton biomass of type $u \in \mathcal{U}$ to 1 BBtu	BBtu/ton
$d_{j,t}$	Demand of biomass from the R.C.I. sectors in state $j \in \mathcal{J}$ in year $t \in \mathcal{T}$	BBtu
$p_{u,j,t}$	Availability of biomass type $u \in \mathcal{U}$ in state $j \in \mathcal{J}$ in year $t \in \mathcal{T}$	ton
p_{v,j,t_0}	Capacity of renewable electricity generation of type $v \in \mathcal{V}$ in state $j \in \mathcal{J}$ in year t_0, which is one year before the first year in the modeling horizon	MW
p_{j,t_0}^{F}	Capacity of cellulosic biofuel production in state $j \in \mathcal{J}$ in year t_0	gallon
α_v	Capacity factor of renewable energy	unitless
$M_{v,j,t}$	Maximum level of new investment in renewable energy facilities in state $j \in \mathcal{J}$ in year $t \in \mathcal{T}$	MW
$M_{j,t}^{\mathrm{F}}$	Maximum level of new investment in cellulosic biofuel facilities in state $j \in \mathcal{J}$ in year $t \in \mathcal{T}$	gallon
$q_{v,j,k}$	Indicator of whether ($q_{v,j,k} = 1$) or not ($q_{v,j,k} = 0$) renewable energy type $v \in \mathcal{V}$ is included in the definition of RPS tier $k \in \mathcal{K}_j$ by state $j \in \mathcal{J}^{\mathrm{RPS}}$	unitless
$e_{j,t}$	Annual electricity consumption projection in state $j \in \mathcal{J}$ in year $t \in \mathcal{T}$	MWh
$\eta_{j,t,k}$	RPS requirements or goals of tier $k \in \mathcal{K}_j$ in state $j \in \mathcal{J}^{\mathrm{RPS}}$ in year $t \in \mathcal{T}$	unitless
θ_t	RFS2 requirements in year $t \in \mathcal{T}$	gallon

Decision variables

Notation	Definition	Unit
ζ	Net present value of total profit throughout the modeling horizon	\$

$x_{u,j,t}$	Biomass production of type $u \in \mathcal{U}$ in state $j \in \mathcal{J}$ in year $t \in \mathcal{T}$	ton
$x_{v,j,t}$	Renewable electricity generation of type $v \in \mathcal{V}$ in state $j \in \mathcal{J}$ in year $t \in \mathcal{T}$	MWh
$x^{\mathrm{F}}_{j,t}$	Cellulosic biofuel production in state $j \in \mathcal{J}$ in year $t \in \mathcal{T}$	gallon
$y_{u,i,j,t}$	Amount of biomass transportation of type $u \in \mathcal{U}$ from state $i \in \mathcal{J}$ to $j \in \mathcal{J}$ in year $t \in \mathcal{T}$	ton
$z_{v,j,t}$	New capacity of renewable electricity generation of type $v \in \mathcal{V}$ in state $j \in \mathcal{J}$ in year $t \in \mathcal{T}$	MW
$z^{\mathrm{F}}_{j,t}$	New capacity of cellulosic biofuel production in state $j \in \mathcal{J}$ in year $t \in \mathcal{T}$	gallon/year
$p_{v,j,t}$	Renewable electricity generation capacity of type $v \in \mathcal{V}$ in state $j \in \mathcal{J}$ in year $t \in \mathcal{T}$	MW
$p^{\mathrm{F}}_{j,t}$	Cellulosic biofuel production capacity in state $j \in \mathcal{J}$ in year $t \in \mathcal{T}$	gallon/year
$s_{j,t,k}$	Renewable electricity generation shortfall of RPS requirements for tier $k \in \mathcal{K}_j$ in state $j \in \mathcal{J}^{\mathrm{RPS}}$ in year $t \in \mathcal{T}$	MWh
s^{F}_{t}	Cellulosic biofuel production shortfall of RFS2 requirements in year $t \in \mathcal{T}$	gallon

2.6 Appendix B

Sets

Notation	Data or data source
$\mathcal{J}$	Data from USA.gov (2013) were used.
$\mathcal{J}^{\mathrm{RPS}}$	Data from DSIRE (2013b) were used.
$\mathcal{K}_j$	Data from DSIRE (2013b) were used.
$\mathcal{T}$	$\mathcal{T} = \{2013, ..., 2035\}$.
$\mathcal{U}$	$\mathcal{U}$ = {agricultural residues, energy crops, forestry residues, urban wood waste/mill residues}, as defined in Haq and Easterly (2006).
$\mathcal{V}$	$\mathcal{V} = \{\text{wind}, \text{geothermal}, \text{solar}, \text{biomass}\}$.

Parameters

Notation	Data or data source
$c_{u,j,t}$	We assume \$96/ton for all types of biomass for all states and a 3.5% annual increase (based on information obtained from personal contact with biofuel companies and research experience).

$\pi_{i,j,t}$	Average transportation cost is obtained from Brechbill and Tyner (2008), and assumed to be \$0.5/(ton mile) for all types of biomass, all states, and all years.
$c_{v,j,t}/f_{v,j,t}$	Average costs from AEO (2012) were used for all states and all years.
$c^{\mathrm{F}}_{j,t}$	Average costs from Tables 14 and 17 of Wright et al. (2010) were used for all states and all years.
$l_{v,j,t}$	Average levelized capital costs in \$/MWh from AEO (2012) were converted to \$/MW using average capacity factor and then used for all states and all years.
$l^{\mathrm{F}}_{j,t}$	Average costs from Tables 13 and 16 of Wright et al. (2010) were used for all states and all years.
$\beta_{j,t}$	Average wholesale electricity prices for all states were obtained from FERC (2013), and the growth rate was estimated from the U.S. average end-use electricity price projection from EIA (2013).
β^{F}_{t}	Motor gasoline prices from AEO (2012) was used as an estimate of the average biofuel price.
$\varphi_{v,t}$	Current values of production tax credits were used with 2029 as the expiration date for all types of renewable energy generation. The current production tax credit policy applies to facilities that begin construction before December 31, 2013, and the credits generally last for 10 years after the facility was placed in service [DSIRE (2013c)]. The expiration dates of 2023 is used for the pessimistic scenario assuming no extension of these credits.
φ^{F}_{t}	Data from AFDC (2013b) were used. The expiration date is assumed to be December of 2022.
$\lambda_{v,t}$	Data from DSIRE (2013a) were used. The expiration date is assumed to be December of 2025 for W.G.S. and biopower.
λ^{F}_{t}	A 15% discount of investment cost was used as the investment tax credit for biofuel production. The expiration date is assumed to be December of 2022. It is stated in AFDC (2013a) that "a second generation biofuel production plant placed into service between December 20, 2006, and December 31, 2013, may be eligible for an additional depreciation tax deduction allowance equal to 50% of the adjusted basis of the property."
$\mu_{j,t,k}$	Data from DSIRE (2013d) were used. Details are also summarized in Table 3 of Cory and Swezey (2007). We made reasonable assumptions for states with non-binding goals or with unclear definitions of penalty, such as using the average of other states' penalties with certain discounts.
μ^{F}_{t}	Assumed to be \$1/gallon for all years, which is in vicinity to recent Renewable Identification Number prices for ethanol.

r	Assumed to be 3.5%.
ρ_u	Average values from Table A-2 of Boundy et al. (2011) were used.
$d_{j,t}$	Data were estimated using historical demand data from the R.C.I. sectors from SEDS (2013) multiplied by the projected growth rate for nine regions in the U.S. Reasonable assumptions were made to assign all states to those regions.
$p_{u,j,t}$	Data from Milbrandt (2005) were used.
p_{v,j,t_0}	Existing capacity for wind, geothermal, solar, and biopower plants were from EERE (2013); GEA (2012); Sherwood (2012), and EIA (2012), respectively.
p^{F}_{j,t_0}	Data from RFA (2012) were used.
α_v	Average annual capacity factors from AEO (2012) were used.
$M_{v,j,t}$	These data were difficult to estimate since all states have their unique strengths and limitations in manufacturing capability, resource (material, labor, funds, etc.) availability, and legislative environment. Our estimates were based on a careful review of all 50 states' and Washington D.C.' resources availability, existing capacity, and historical growth rate.
$M^{\mathrm{F}}_{j,t}$	There is no cellulosic biofuel facilities operating in the U.S. However, according to Brown and Brown (2013), nine commercial-scale facilities in eight states are expected to be in operation by 2014. For these eight states, we set the investment limit as twice of the expected capacity by 2014; for other states, the limit is assumed to be 30 million gallons/year.
$q_{v,j,k}$	Data from DSIRE (2013d) were used.
$e_{j,t}$	Similar to $d_{j,t}$, annual electricity consumption for each state was projected through 2035.
$\eta_{j,t,k}$	Data from DSIRE (2013d) were used. RPS requirements for almost all states were defined in percentages of total electricity consumption. Two exceptions are Iowa and Texas, which mandated renewable electricity generation capacity (MW). Appropriate adjustments were made for these two states.
θ_t	Data from Schnepf and Yacobucci (2010) were used.

BIBLIOGRAPHY

AEO (2012). Annual energy outlook 2012 with projections to 2035. Technical report, U.S. Energy Information Administration (EIA).

AFDC (accessed in 2013a). Second Generation Biofuel Plant Depreciation Deduction Allowance. http://www.afdc.energy.gov/laws/law/US/413.

AFDC (accessed in 2013b). Second Generation Biofuel Producer Tax Credit. http://www.afdc.energy.gov/laws/law/US/10515.

Boundy, B., Diegel, S. W., Wright, L., and C., D. S. (2011). Biomass energy data book: Edition 4. Technical report, U.S. DOE, Energy Efficiency and Renewable Energy (EERE).

Brechbill, S. C. and Tyner, W. E. (2008). The economics of biomass collection, transportation, and supply to indiana cellulosic and electric utility facilities. Technical report.

Brown, T. R. and Brown, R. C. (2013). A review of cellulosic biofuel commercial-scale projects in the United States. *Biofuels, Bioproducts and Biorefining*.

Carley, S. (2009). State renewable energy electricity policies: An empirical evaluation of effectiveness. *Energy Policy*, 37(8):3071–3081.

Cory, K. S. and Swezey, B. G. (2007). Renewable Portfolio Standards in the States: Balancing goals and rules. *The Electricity Journal*, 20(4):21–32.

Dassanayake, G. D. M. and Kumar, A. (2012). Techno-economic assessment of triticale straw for power generation. *Applied Energy*, 98:236–245.

DSIRE (accessed in 2013a). Business energy investment tax credit (itc). http://www.dsireusa.org/incentives/incentive.cfm?Incentive_Code=US02F.

DSIRE (accessed in 2013b). Database of state incentives for renewables and efficiency. http://www.dsireusa.org.

DSIRE (accessed in 2013c). Renewable Electricity Production Tax Credit (PTC). http://www.dsireusa.org/incentives/incentive.cfm?Incentive_Code=US13F.

DSIRE (accessed in 2013d). RPS data. http://www.dsireusa.org/rpsdata/index.cfm.

EERE (accessed in 2013). Installed wind capacity. http://www.windpoweringamerica.gov/wind_installed_capacity.asp.

EIA (2012). State renewable electricity profiles 2010. Technical report, U.S. Energy Information Administration (EIA).

EIA (accessed in 2013). Electricity supply, disposition, prices, and emissions, reference case. http://www.eia.gov/oiaf/aeo/tablebrowser/#release=AEO2012&subject=3-AEO2012&table=8-AEO2012&region=0-0&cases=ref2012-d020112c.

FERC (accessed in 2013). Electric power markets: National overview. http://www.ferc.gov/market-oversight/mkt-electric/overview.asp.

GEA (2012). Annual u.s. geothermal power production and development report. Technical report, Geothermal Energy Association (GEA). http://www.geo-energy.org/.

Haq, Z. and Easterly, J. L. (2006). Agricultural residue availability in the united states. *Applied Biochemistry and Biotechnology*, 129-132:3–21.

Huang, H., Khanna, M., Onal, H., and Chen, X. (2013). Stacking low carbon policies on the Renewable Fuels Standard: Economic and greenhouse gas implications. *Energy Policy*, 56:5–15.

Jeffers, R. F., Jacobson, J. J., and Searcy, E. M. (2013). Dynamic analysis of policy drivers for bioenergy commodity markets. *Energy Policy*, 52:249–263.

Johnson, S. D. and Moyer, E. J. (2012). Feasibility of U.S. Renewable Portfolio Standards under cost caps and case study for Illinois. *Energy Policy*, 49:499–514.

Menza, F. C. and Vachon, S. (2006). The effectiveness of different policy regimes for promoting wind power: Experiences from the states. *Energy Policy*, 34(14):1786–1796.

Milbrandt, A. (2005). A geographic perspective on the current biomass resource availability in the United States. Technical report, National Renewable Energy Laboratory (NREL).

Palmer, K. and Burtraw, D. (2005). Cost-effectiveness of renewable electricity policies. *Energy Economics*, 27(6):873–894.

RFA (2012). 2012 ethanol industry outlook. Technical report, Renewable Fuels Association (RFA).

Schnepf, R. and Yacobucci, B. D. (2010). Renewable fuel standard (rfs): Overview and issues. Technical report, Congressional Research Service.

SEDS (accessed in 2013). State energy data system (seds): 1960-2010. http://www.eia.gov/state/seds/seds-data-complete.cfm?sid=US#Consumption.

Sherwood, L. (2012). U.s. solar market trends 2011. Technical report, Interstate Renewable Energy Council, Inc.

USA.gov (accessed in 2013). 50 states and the District of Columbia. http://www.usa.gov/Agencies/State-and-Territories.shtml.

Wiser, R., Namovicz, C., Gielecki, M., and Smith, R. (2007). The experience with Renewable Portfolio Standards in the United States. *The Electricity Journal*, 20(4):8–20.

Wright, M. M., Satrio, J. A., Brown, R. C., Daugaard, D. E., and Hsu, D. D. (2010). Techno-economic analysis of biomass fast pyrolysis to transportation fuels. Technical report, National Renewable Energy Laboratory (NREL).

Yin, H. and Powers, N. (2010). Do state Renewable Portfolio Standards promote in-state renewable generation? *Energy Policy*, 38(2):1140–1149.

CHAPTER 3. A TRI-LEVEL OPTIMIZATION MODEL FOR INVENTORY CONTROL WITH UNCERTAIN DEMAND AND LEAD TIME

A paper submitted to European Journal of Operational Research

Mohammad Rahdar, Lizhi Wang, and Guiping Hu

Abstract

We propose an inventory control model for an uncapacitated warehouse in a manufacturing facility under demand and lead time uncertainty. The objective is to make strategic ordering decisions to minimize the total system cost. We introduce a two-stage tri-level optimization model with a rolling planning horizon to address the uncertain demand and lead time regardless of their underlying distributions. In addition, an exact algorithm is designed to solve the model. We compare this model with three deterministic models in a case study. Our computational results suggest that the performances of deterministic models are either consistently inferior or highly sensitive to cost parameters (such as holding cost and shortage cost), whereas the new tri-level optimization model almost always results in the lowest total cost in all parameter settings.

3.1 Introduction

Uncertainty along a supply chain network is ubiquitous; it may arise for the arrival of raw materials or it may appear over customer demands. Since the stakeholders along the supply chain are interconnected, inventory systems are often complicated concerning uncertainty and variability. Several studies Davis (1993); Ho et al. (2005); Wang and Shu (2005); Li and Schulze (2011) have mentioned that there are typically three sources of uncertainty in a supply chain:

suppliers, manufacturing, and customers. Supplier uncertainty leads to variability in lead time and customer uncertainty appears in order time or quantity, both of which would cause unexpected costs.

Most studies on inventory control systems focused on addressing uncertainty from either the demand or supply side. Axsäter (2003) and Seifbarghy and Jokar (2006) proposed a model with a central warehouse and several retailers to estimate the optimal reorder point when the demand was uncertain. Routroy and Kodali (2005) studied a supply chain including a manufacturer, a distributor, and a retailer with an uncertain demand to minimize the total system cost. In addition, Wang (2013) considered a two-level supply chain with one warehouse and multiple retailers and assumed that retailers faced independent Poisson demand processes. Moreover, in the model proposed by Muriana (2016), demand rate for perishable products was a random variable following a normal distribution. On the other hand, significant research has been also done to address the uncertainty of lead time. Sajadieh et al. (2009) proposed a model to minimize the total cost of an integrated vendor-buyer supply chain when the lead time is stochastic. Furthermore, Hoque (2013) assumed that the lead time was an independent random variable from a normal distribution. Maiti et al. (2009) developed an inventory model where the lead time was a random variable which followed either normal or exponential distributions. Another approach of considering lead time was described by Dey et al. (2008), who developed a finite time horizon inventory model with interval-valued lead time. Few studies have been devoted to addressing uncertainty from both suppliers and customers. However, both sources of uncertainty and their interactions could have convoluted implications to the entire supply chain. In this paper, we propose a new inventory control model that takes into account both lead time uncertainty and demand variability.

It has been shown that if the probabilistic description of randomness is available, stochastic programming is an effective tool to address uncertainty, but this information is not always available in real applications (Bertsimas and Thiele, 2006; Unlu and Rossetti, 2009). As reported by Pan and Nagi (2010), supply chain models with stochastic parameters can be classified into two main approaches, probabilistic approach and scenario approach. When there is probability information about uncertain parameters, the parameters can be considered as random variables

in the probabilistic approach. Otherwise, uncertainty can be characterized by defining a set of scenarios, which represents a number of potential future states (Pan and Nagi, 2010). This paper presents a novel method of determining scenarios, and obtaining optimal solution under the worst-case scenario.

We introduce an inventory control model for a warehouse in a manufacturing facility, which serves the demand for a single item. The goal is to define the order policy to minimize system costs. Demand and lead time are uncertain parameters, and the probability distributions are unknown. The only available information is that uncertain parameters are independent random variables that can take some values from their intervals. In addition, the shortage is allowed and fully backlogged. The objective is to determine the time and size of orders, such that the total cost, which consists of order, inventory holding, and shortage costs, is minimized. Since uncertain demand is observed in each period and the exact lead time is realized when the order arrives, it is a multi-stage decision-making problem. We approximate it by developing a two-stage tri-level optimization model to reduce the curse of dimensionality. This simplified model is solved in a rolling horizon framework. Under this approach, the first stage decisions are implemented; then, the next planning horizon is planned with updated information (Sahin et al., 2013).

The remainder of the paper is organized as follows: in Section 3.2, detailed problem formulation is discussed. Section 3.3 is devoted to algorithm development. Section 3.4 presents the experimental results and sensitivity analysis. Finally, the conclusion with a summary is reported in Section 3.5.

3.2 Model formulation

3.2.1 Problem statement

We consider an uncapacitated warehouse for a single item in a manufacturing facility. The demand and lead time are both uncertain. Decisions are made over an indefinite discrete time period to minimize the order, inventory, and shortage costs. We assume that shortage is fully backlogged, and demand and orders come at the beginning of the decision period, and the

manager has full information about the demand, current inventory/shortage, and order arrival status to make an order decision for that period.

For modeling purposes, we label the current period as period 1 and we impose a finite planning horizon $\{1, 2, \ldots, T\}$. The solution from this model can be applied in a rolling horizon manner, in which the model is solved in each decision period with updated information and only the order decision for the current period is actually executed. This process is illustrated in Figure 3.1. The decision making model $P(\tau)$ has a planning horizon from period τ to $\tau+T-1$. After solving the decision-making model $P(\tau)$, and determining the order policy, we divide the decision of the planning horizon into two parts: the decision of the first period, $\{\tau\}$, and the decision of the second period and afterward, $\{\tau+1, \ldots, \tau+T-1\}$. Order policy of period τ is implemented and τ is increased by 1, the initial parameters of the next planning horizon are updated, and the model is run again. Therefore, the decision of periods $\{\tau+1, \ldots, \tau+T-1\}$ may reschedule in the next planning horizon. Solid lines in Figure 3.1 indicate the fixed decisions.

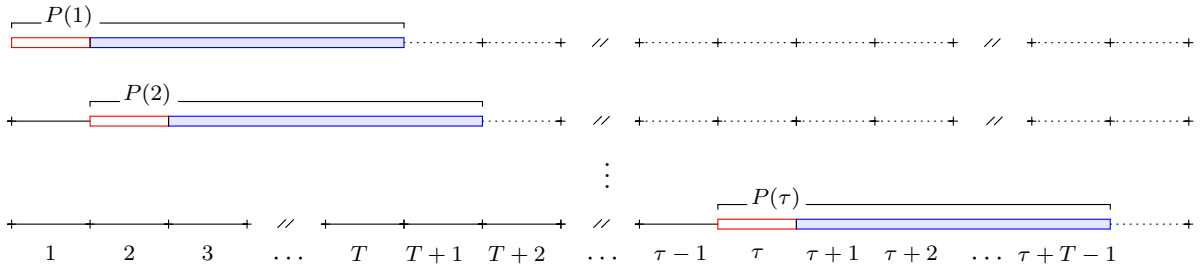


Figure 3.1: Rolling planning horizon approach

The fidelity of the aforementioned planning model largely depends on the planning horizon parameter T. From a computational tractability perspective, due to the well-known curse of dimensionality (Defourny et al., 2011), multi-stage decision-making models with $T \geq 3$ are notoriously hard to solve. From a practical perspective, however, models with such a small planning horizon are systemically shortsighted and may yield solutions that are too myopic to be practically useful. Our proposed approach is a tri-level optimization model that represents a compromise between these two competing perspectives. In the remainder of the section, we first give the deterministic version of the planning model in Section 3.2.2 for benchmark purposes and then introduce the tri-level optimization model in Section 3.2.3.

3.2.2 Deterministic model

Consider a simplified version of the inventory control model where the demand and lead time in all periods are assumed to be constant and known. As such, the multi-stage decision-making problem reduces to a deterministic single stage optimization model.

Table 3.1: Notation in the deterministic model

Decision variables	
$q_t \in \mathbb{Z}^+$	Number of batches ordered in period t, $\forall t \in \{1, 2, \ldots, T\}$
$I_t \in \mathbb{Z}^+$	Inventory level in period t, $\forall t \in \{1, 2, \ldots, T\}$
$g_t \in \mathbb{Z}^+$	Shortage amount in period t, $\forall t \in \{1, 2, \ldots, T\}$
$v_t \in \mathbb{B}$	Indicating whether an order is placed in period t ($v_t = 1$) or not ($v_t = 0$), $\forall t \in \{1, 2, \ldots, T\}$
Parameters	
c	Variable order cost
f	Fixed order cost
h	Inventory holding cost
p	Shortage cost
T	Number of periods in the planning horizon
M	A sufficiently large positive number (big-M)
μ	Order batch size
I_0	Initial inventory level at the beginning the planning horizon
K	Number of periods before the planning horizon with orders on the way
q_k	Number of batches ordered in period k, $\forall k \in \{1 - K, \ldots, -1, 0\}$ before the planning horizon
$\hat{d}_t$	Assumed demand of period t, $\forall t \in \{1, \ldots, T\}$
$\hat{\delta}_{k,t}$	Assumed order arrival status, indicating whether ($\hat{\delta}_{k,t} = 1$) or not ($\hat{\delta}_{k,t} = 0$) the order made in period k arrives by period t, $\forall k \in \{1 - K, \ldots, t - 1\}, \forall t \in \{k + 1, \ldots, T\}$

Table 3.1 includes the notations used in formulating the deterministic model. It is worth noting that the random lead time is represented by a set of binary parameters $\delta_{k,t}, \forall k, t$, indicating whether or not the order made in period k arrives by period t. For example, if the lead time of an order made in period 3 is 4, then $\delta_{3,4} = \delta_{3,5} = \delta_{3,6} = 0$ and $\delta_{3,t} = 1, \forall t \in \{7, 8, \ldots T\}$.

The deterministic inventory control model is given in (3.1a)-(3.1d). The objective of the model is to minimize the total cost over the planning horizon. The four cost terms in (3.1a) are the variable order cost, fixed order cost, inventory holding cost, and shortage cost, respectively. Equation (3.1b) calculates the inventory level at the end of period t. The four terms on the

right-hand-side of Constraint (3.1b) are, respectively, the initial inventory at period 0, the total amount of ordered items that arrive by period t, the amount of shortage at period t, and the total amount of demand that is served between periods 1 and t. Constraint (3.1c) ensures that a fixed order cost is incurred if at least one item is ordered in that period. The supports of the decision variables are defined in Constraint (3.1d).

$$\min \quad \zeta = c\mu \sum_{t=1}^{T} q_t + f \sum_{t=1}^{T} v_t + h \sum_{t=1}^{T} I_t + p \sum_{t=1}^{T} g_t \tag{3.1a}$$

$$\text{s.t.} \quad I_t = I_0 + \sum_{k=1-K}^{t-1} \mu q_k \hat{\delta}_{k,t} + g_t - \sum_{i=1}^{t} \hat{d}_i \qquad t \in \{1, 2, \ldots, T\} \tag{3.1b}$$

$$q_t \leq M v_t \qquad t \in \{1, 2, \ldots, T\} \tag{3.1c}$$

$$q_t, I_t, g_t \in \mathbb{Z}^+; v_t \in \mathbb{B} \qquad t \in \{1, 2, \ldots, T\} \tag{3.1d}$$

3.2.3 Tri-level optimization model

Relaxing the simplifying assumptions on perfect information of demand and lead time results in a multi-stage decision-making problem, in which uncertain demand is observed in each period but the exact lead time is not realized until when the order arrives. We propose a two-stage tri-level optimization model to approximate the multi-stage decision-making problem and to alleviate its curse of dimensionality. The first stage refers to the first period of the planning horizon, whereas all the remaining periods are aggregated into the second stage; a similar modeling approach has been suggested by Defourny et al. (2011). As such, after the first stage decision has been made, all uncertain parameters for period 2 and beyond are assumed to be observable, and thus, the second stage becomes a deterministic problem. We further assume that the first stage will take a pessimistic view of uncertainty and anticipate the worst-case scenario for the second stage. Therefore, the two-stage decision-making model is formulated as a tri-level optimization model, in which the upper-level makes the first stage decision, the middle-level identifies the worst-case scenario given the first stage decision, and the lower-level makes the second stage decision given the first stage decision and under the worst-case scenario. This simplified model may become more appropriate in a rolling horizon framework (Beaudin and Zareipour, 2015), in which the tri-level model is solved in every period with updated

information, but only the first stage decisions are implemented. The determination of the first stage decisions is balanced between a pessimistic anticipation of the worst-case scenario and an optimistic assumption of perfect information throughout the rest of the planning horizon.

The tri-level optimization model is developed using notations defined in Table 3.2. The assumption is that demands and lead times are uncertain, but we know the lower and upper bounds of these uncertain parameters, which are time dependent and independent of each other. It should be noted that for $t \in \{2, \ldots, T\}$, demand and order arrival status were defined as parameters in Table 3.1, but they become the middle-level decision variables in the tri-level optimization model.

Table 3.2: Notation in the tri-level model

Decision variables for the upper-level	
$q_1 \in \mathbb{Z}^+$	Number of batches ordered in period 1
$I_1 \in \mathbb{Z}^+$	Inventory level in period 1
$g_1 \in \mathbb{Z}^+$	Shortage amount in period 1
$v_1 \in \mathbb{B}$	Indicating whether an order is placed in period 1 ($v_1 = 1$) or not ($v_1 = 0$)
x	Aggregated upper-level decision variables, $x = [q_1, I_1, g_1, v_1]^\top$
Decision variables for the middle-level	
$d_t \in \mathbb{Z}^+$	Demand of period $t, \forall t \in \{2, \ldots, T\}$
$\delta_{k,t} \in \mathbb{B}$	Order arrival status, indicating whether ($\delta_{k,t} = 1$) or not ($\delta_{k,t} = 0$) the order made in period $k, \forall k \in \{1-K, \ldots, T-1\}$ arrives by period $t, \forall t \in \{k+1, \ldots, T\}$
y	Aggregated middle-level decision variables, $y = [d_2, \ldots, d_T, \delta_{1-K,2}, \ldots, \delta_{1,2}, \delta_{1-K,3}, \ldots, \delta_{2,3}, \ldots, \delta_{T-1,T}]^\top$
Decision variables for the lower-level	
$q_t \in \mathbb{Z}^+$	Number of batches ordered in period t, $\forall t \in \{2, 3, \ldots, T\}$
$I_t \in \mathbb{Z}^+$	Inventory level in period t, $\forall t \in \{2, 3, \ldots, T\}$
$g_t \in \mathbb{Z}^+$	Shortage amount in period t, $\forall t \in \{2, 3, \ldots, T\}$
$v_t \in \mathbb{B}$	Indicating whether an order is placed in period t ($v_t = 1$) or not ($v_t = 0$), $\forall t \in \{2, 3, \ldots, T\}$
z	Aggregated lower-level decision variables $z = [q_2, \ldots, q_T, I_2, \ldots, I_T, g_2, \ldots, g_T, v_2, \ldots, v_T]^\top$

Parameters

c	Variable order cost
f	Fixed order cost
h	Inventory holding cost
p	Shortage cost
T	Number of periods in the planning horizon
M	A sufficiently large positive number (big-M)
μ	Order batch size
I_0	Initial inventory level at the beginning the planning horizon
K	Number of periods before the planning horizon with orders on the way
q_k	Number of batches ordered in period k, $\forall k \in \{1-K, \ldots, -1, 0\}$ before the planning horizon
l_t^{D}	Lower bound of demand in period $t, \forall t \in \{1, 2, \ldots, T\}$
u_t^{D}	Upper bound of demand in period $t, \forall t \in \{1, 2, \ldots, T\}$
l_k^{L}	Lower bound of lead time for the order placed in period $k, \forall k \in \{1-K, \ldots, T-1\}$
u_k^{L}	Upper bound of lead time for the order placed in period $k, \forall k \in \{1-K, \ldots, T-1\}$
$\tilde{d}_1$	Observed demand of period 1
$\tilde{\delta}_{k,1}$	Observed order arrival status, indicating whether ($\tilde{\delta}_{k,1} = 1$) or not ($\tilde{\delta}_{k,1} = 0$) the order made in period $k, \forall k \in \{1-K, \ldots, -1, 0\}$ arrives by period 1.
c_1	Aggregated objective function coefficients of the first stage decisions, $c_1 = [c\mu, f, h, p]^\top$
c_2	Aggregated objective function coefficients of the second stage decisions, $c_2 = [c\mu, \ldots, c\mu, f, \ldots, f, h, \ldots, h, p, \ldots, p]^\top$

Using the notations of aggregated decision variables and parameters, we formulate the tri-level optimization model as follows.

$$\min_{x \in \mathcal{X}} \left\{ c_1^\top x + \max_{y \in \mathcal{Y}(x)} \left\{ \min_{z \in \mathcal{Z}(x,y)} \; c_2^\top z \right\} \right\}. \tag{3.2}$$

Here, the lower-level solves a deterministic problem, $\min_{z \in \mathcal{Z}(x,y)} c_2^\top z$, to minimize the total cost for periods 2 to T given the first stage order decision, x, made at the upper-level and the

worst-case scenario, y, identified by the middle-level. The feasible set $\mathcal{Z}(x,y)$ is defined as

$$\mathcal{Z}(x,y) = \left\{ \begin{array}{lll} z: & I_t = I_0 + \sum_{k=1-K}^{t-1} \mu q_k \delta_{k,t} + g_t - \sum_{i=1}^{t} d_i & \forall t \in \{2,3,\ldots,T\} \\ & q_t \leq M v_t & \forall t \in \{2,3,\ldots,T\} \\ & q_t, I_t, g_t \in \mathbb{Z}^+, v_t \in \mathbb{B} & \forall t \in \{2,3,\ldots,T\} \end{array} \right\}. \quad (3.3)$$

Notice that the term $\sum_{k=1-K}^{t-1} \mu q_k \delta_{k,t}$ is nonlinear, since both q_k and $\delta_{k,t}$ are part of decision variables z and y, respectively. We will linearize this term in Section 3.3.

The middle-level observes the order decision, x, made at the upper-level and solves a bilevel optimization model, $\max_{y \in \mathcal{Y}(x)} \left\{ \min_{z \in \mathcal{Z}(x,y)} \; c_2^\top z \right\}$, to identify the worst-case scenario, anticipating the response of the lower-level. The feasible set $\mathcal{Y}(x)$ is defined as

$$\mathcal{Y}(x) = \left\{ \begin{array}{lll} y: & \tilde{\delta}_{k,1} \leq \delta_{k,2} & \forall k \in \{1-K,\ldots,0\} \\ & \delta_{k,t} \leq \delta_{k,t+1} & \forall k \in \{1-K,\ldots,T-2\}, \forall t \in \{\max\{k+1,2\},\ldots,T-1\} \\ & l_t^{\mathrm{D}} \leq d_t \leq u_t^{\mathrm{D}} & \forall t \in \{2,3,\ldots,T\} \\ & l_k^{\mathrm{L}} \leq 1 + \sum_{t=k+1}^{T} (1-\delta_{k,t}) \leq u_k^{\mathrm{L}} & \forall k \in \{1-K,\ldots,T-1\} \\ & d_t \in \mathbb{Z}^+ & \forall t \in \{2,3,\ldots,T\} \\ & \delta_{k,t} \in \mathbb{B} & \forall k \in \{1-K,\ldots,T-1\}, \forall t \in \{\max\{k+1,2\},\ldots,T\} \end{array} \right\}.$$

The first and second constraints ensure that once an order arrives in period t, all subsequent status variables must be set as $\delta_{k,\tau} = 1, \forall \tau \geq t$. The third and fourth constraints set the lower and upper bounds for demand and lead time in the second stage periods, respectively.

The upper-level solves the tri-level optimization model (3.2), which minimizes the combined cost terms for period 1, $c_1^\top x$, and for the rest of the planning horizon, $c_2^\top z$, anticipating the response from the middle and lower levels. The feasible set $\mathcal{X}$ is defined as

$$\mathcal{X} = \left\{ \begin{array}{lll} x: & I_1 = I_0 + \sum_{k=1-K}^{0} \mu q_k \tilde{\delta}_{k,1} + g_1 - \tilde{d}_1 \\ & q_1 \leq M v_1 \\ & q_1, I_1, g_1 \in \mathbb{Z}^+, v_1 \in \mathbb{B} \end{array} \right\}.$$

3.3 Algorithm design

We define $\tilde{\mathcal{Y}} = \bigcup_{x \in \mathcal{X}} \mathcal{Y}(x)$ and let $\{y^i : \forall i \in \mathcal{I}\}$ denote all the elements in set $\tilde{\mathcal{Y}}$, where $\mathcal{I}$ is the set of superscripts for y^i with $|\mathcal{I}| = |\tilde{\mathcal{Y}}|$. Then model (3.2) is equivalent to

$$\min_{x,z,\xi} \left\{ c_1^\top x + \xi : x \in \mathcal{X}; \xi \geq c_2^\top z^i, z^i \in \mathcal{Z}(x,y^i), \forall i \in \mathcal{I} \right\}. \quad (3.4)$$

Here, instead of treating the worst-case scenario y as a decision variable for the middle-level, we consider all possible scenarios of $y^i, \forall i \in \mathcal{I}$ as given parameters and define a response variable z^i for each possible scenario y^i. The constraints $\xi \geq c_2^\top z^i, \forall i \in \mathcal{I}$ and the objective function $c_1^\top x + \xi$ ensure that only the worst-case scenario cost is being minimized. As such, the middle-level is eliminated, and the upper and lower levels merge into one single level optimization model (3.4). This reformulation is challenged by the potentially enormous number of additional decision variables z^i and constraints, which may make it computationally intractable.

We propose an exact algorithm for model (3.2) by using the reformulation (3.4) and overcoming the challenges with its dimensions. The steps of the algorithm are described in Algorithm 3.1. The idea is to solve model (3.4) with a small subset $\hat{\mathcal{Y}} \subseteq \tilde{\mathcal{Y}}$ of scenarios, which is a relaxation of (3.4), and iteratively add new scenarios. Such scenarios are generated in line 10 by solving the middle and lower levels with fixed upper-level decisions from the relaxation solution. The resulting bilevel model either confirms the optimality of the upper-level decision or yields a worst-case scenario that will be included in $\hat{\mathcal{Y}}$ in the next iteration.

Algorithm 3.1 Algorithm of solving the tri-level model (3.2)

1: Inputs: $\mathcal{X}, \tilde{\mathcal{Y}}$, and $\mathcal{Z}(x,y), \forall x \in \mathcal{X}, y \in \tilde{\mathcal{Y}}$
2: Initialize $(x^*, y^*, z^*) = \emptyset, \zeta^{\mathrm{L}} = -\infty, \zeta^{\mathrm{U}} = \infty$
3: Identify a set $\hat{\mathcal{Y}}$ such that $\emptyset \subset \hat{\mathcal{Y}} \subseteq \tilde{\mathcal{Y}}$ and define $\hat{\mathcal{I}} = \{i : \forall y^i \in \hat{\mathcal{Y}}\}$
4: **while** $\zeta^{\mathrm{L}} < \zeta^{\mathrm{U}}$ **do**
5: Solve the following Master problem
$\mathcal{M}(\hat{\mathcal{I}})$: $\min_{x,z,\xi} \left\{ c_1^\top x + \xi : x \in \mathcal{X}; \xi \geq c_2^\top z^i, z^i \in \mathcal{Z}(x, y^i), \forall i \in \hat{\mathcal{I}} \right\}$
6: **if** infeasible **then**
7: Return model (3.2) is infeasible
8: **else**
9: Let $(\hat{x}, \hat{\xi})$ denote the corresponding components of an optimal solution
10: Solve the following Subproblem $\mathcal{S}(\hat{x})$: $\max\limits_{y \in \mathcal{Y}(\hat{x})} \left\{ \min\limits_{z \in \mathcal{Z}(\hat{x}, y)} c_2^\top z \right\}$ and let $(\hat{y}, \hat{z})$ denote an optimal solution
11: Update $\zeta^{\mathrm{L}} \leftarrow c_1^\top \hat{x} + \hat{\xi}$, $\zeta^{\mathrm{U}} \leftarrow \max\{\zeta^{\mathrm{U}}, c_1^\top \hat{x} + c_2^\top \hat{z}\}$, $\hat{\mathcal{Y}} \leftarrow \hat{\mathcal{Y}} \cup \{\hat{y}\}$, and $\hat{\mathcal{I}} \leftarrow \{i : \forall y^i \in \hat{\mathcal{Y}}\}$
12: **end if**
13: **end while**
14: Return $x^* = \hat{x}, y^* = \hat{y}, z^* = \hat{z}$

Since y and z are treated as variables in the Subproblem $\mathcal{S}(\hat{x})$, the multiplication of q_k and $\tilde{\delta}_{k,t}$ introduces nonlinearity to the set $\mathcal{Z}(\hat{x}, y)$, which was defined in (3.3). To linearize the set $\mathcal{Z}(\hat{x}, y)$, we introduce new variables $u_{k,t} = q_k \tilde{\delta}_{k,t}, \forall k \in \{2, ..., T-1\}, t \in \{k+1, ..., T\}$. Accordingly, we add four new sets of constraints. Variable $u_{k,t}$ is equal to q_k if the order made in period k arrives by period t; otherwise, it is zero. The linearized set $\mathcal{Z}(\hat{x}, y)$, denoted as $\tilde{\mathcal{Z}}(\hat{x}, y)$, is defined as follows.

$$\tilde{\mathcal{Z}}(\hat{x}, y) = \left\{ \begin{array}{lll} z: & I_t = I_0 + \sum\limits_{k=1-K}^{1} \mu q_k \delta_{k,t} + \sum\limits_{k=2}^{t-1} \mu u_{k,t} + g_t - \sum\limits_{i=1}^{t} d_i & t \in \{2, 3, \ldots, T\} \\ & q_t \leq M v_t & t \in \{2, 3, \ldots, T\} \\ & u_{k,t} \geq q_k - M(1 - \delta_{k,t}) & k \in \{2, \ldots, T-1\}, t \in \{k+1, \ldots, T\} \\ & u_{k,t} \leq M \delta_{k,t} & k \in \{2, \ldots, T-1\}, t \in \{k+1, \ldots, T\} \\ & u_{k,t} \leq q_k & k \in \{2, \ldots, T-1\}, t \in \{k+1, \ldots, T\} \\ & u_{k,t} \geq 0 & k \in \{2, \ldots, T-1\}, t \in \{k+1, \ldots, T\} \\ & q_t, I_t, g_t \in \mathbb{Z}^+, v_t \in \mathbb{B} & t \in \{2, 3, \ldots, T\} \end{array} \right\}$$

The resulting Subproblem $\mathcal{S}(\hat{x})$ is a bi-level integer linear programming problem, which can be solved by existing algorithms such as Xu and Wang (2014). The algorithm is able to find the optimal solution to model (3.2) in no more than $(|\mathcal{X}| + 1)$ iterations, which is a finite number since $\mathcal{X}$ is a finite set. For all $i \in \{1, \ldots, |\mathcal{X}| + 1\}$, let $\hat{x}^i$ denote the solution from line 9 in the ith iteration, then there must exist $1 \leq j < k \leq |\mathcal{X}| + 1$ such that $\hat{x}^j = \hat{x}^k$.

3.4 Numerical results

We conducted a simulation experiment to test and compare the performances of the tri-level optimization model and three deterministic models, which we will refer to as Model 1, Model 2, and Model 3. The three deterministic models all use the same formulation (3.1) but with different assumptions about the data. Model 1 uses the 20th percentile of demand and lead time from the simulation samples, Model 2 uses the arithmetic mean, and Model 3 uses the 80th percentile.

3.4.1 Simulation setup

The data used in the simulation study are summarized in Table 3.3. For each of the four models, a total of nine sets of experiments were conducted with all possible combinations of $c = 2$, $h = \{1, 5, 15\}$, and $p = \{4, 10, 30\}$, each of which had 500 repetitions. We generated simulation data for a time horizon of 36 periods $\{-1, 0, ..., 34\}$, of which only 30 periods in the middle $\{1, ..., 30\}$ are used to test the four models and measure their performances. The random lead times $\tilde{L}_k(s), \forall k \in \{-1, 0, \ldots, 34\}, s \in \{1, 2, \ldots, 500\}$ were generated but never used directly in any of the models; rather, they were used to calculate the order arrival statuses $\tilde{\delta}_{k,t}(s)$.

The simulation was carried out in the following manner with 500 random repetitions. We first generated random values for $q_k(s)$ and $\tilde{L}_k(s)$ for the pre-planning period of $k \in \{-1, 0\}$ and then calculated $\hat{\delta}_{-1,0}(s) = \begin{cases} 1 & \text{if } L_{-1} = 1 \\ 0 & \text{otherwise,} \end{cases}$ and $I_0(s) = \mu q_{-1}(s)\hat{\delta}_{-1,0}(s)$. For each of the planning periods $\tau \in \{1, ..., 30\}$, we generated random values for $\hat{d}_\tau(s)$ and $L_\tau(s)$ and updated two other parameters for the four models with $I_0 = I_{\tau-1}$ and $\hat{\delta}_{k,\tau} = \begin{cases} 1 & \text{if } \tau - k \geq L_k(s) \\ 0 & \text{otherwise} \end{cases}, \forall k \in \{\tau - 2, \tau - 1\}$, where $I_{\tau-1}$ is from the optimal solution of planning period $\tau - 1$. We also updated the assumed order arrival status values for the three deterministic models for all $k \in \{\tau - 2, ..., \tau + 3\}$ and $t \in \{\max\{k, \tau\} + 1, ..., \tau + 4\}$ as follows. For Model 1, $\hat{\delta}_{k,t} = \begin{cases} 1 & \text{if } t - k \geq 1 \\ 0 & \text{otherwise} \end{cases}$; for Model 2, $\hat{\delta}_{k,t} = \begin{cases} 1 & \text{if } t - k \geq 1 + (t \bmod 2) \\ 0 & \text{otherwise} \end{cases}$; and for Model 3, $\hat{\delta}_{k,t} = \begin{cases} 1 & \text{if } t - k \geq 2 \\ 0 & \text{otherwise} \end{cases}$. This means that Model 1 and Model 3 assume that the lead time for all orders to be made in the future will be 1 and 2, respectively, and Model 2 assumes that the lead time will be alternatively 1 or 2.

We ran each model 30 times through the simulation experiment from $\tau = 1$ to $\tau = 30$. Order policy of period τ is implemented, the total cost of period τ is saved, and τ is increased by 1 to run the model again. Each box, $P(\tau)$, in Figure 3.2 represents a decision making model for period τ, which has a planning horizon of $\{\tau, \tau + 1, \ldots, \tau + T - 1\}$. The downward arrows into the box represent observed realizations of uncertain demand, $\tilde{d}_\tau$, and order arrival status, $\{\tilde{\delta}_{\tau-K,\tau}, \tilde{\delta}_{\tau-K+1,\tau}, \ldots, \tilde{\delta}_{\tau-1,\tau}\}$. Here, the binary uncertainty parameter $\tilde{\delta}_{k,\tau}$ indicates whether $(\tilde{\delta}_{k,\tau} = 1)$ or not $(\tilde{\delta}_{k,\tau} = 0)$ the items that were ordered in period k arrive in or before period

Table 3.3: Simulation data

c	2
f	0
h	{1,5,15}
p	{4,10,30}
T	5
M	2000
μ	1
K	2
$\hat{d}_t(s)$	Integer randomly generated from a negative binomial distribution with $r = 15$ and $p = 0.3$ for period $t \in \{1, 2, \ldots, 34\}$ and repetition $s \in \{1, 2, \ldots, 500\}$
$\bar{d}_t$	$\bar{d}_t = \frac{r(1-p)}{p} = 35, \forall t \in \{1, 2, \ldots, 34\}$
l_t^{D}	26, the 20th percentile of the negative binomial distribution with parameters $r = 15$ and $p = 0.3$ for period $t \in \{1, 2, \ldots, 34\}$
u_t^{D}	44, the 80th percentile of the negative binomial distribution with parameters $r = 15$ and $p = 0.3$ for period $t \in \{1, 2, \ldots, 34\}$
$\hat{L}_k(s)$	Integer randomly generated from a uniform distribution within $[1, 2]$, representing the lead time for the order placed in period $k, \forall k \in \{-1, 0, \ldots, 34\}$ and repetition $s \in \{1, 2, \ldots, 500\}$
l_k^{L}	1
u_k^{L}	2
$\hat{\delta}_{k,t}(s)$	Integer calculated as $\hat{\delta}_{k,t}(s) = \begin{cases} 1 & \text{if } t - k \geq L_k(s) \\ 0 & \text{otherwise} \end{cases}$, which indicates whether the order placed in period $k \in \{-1, 0, \ldots, 33\}$ arrives by period $t \in \{k + 1, \ldots, 34\}$ for repetition $s \in \{1, 2, \ldots, 500\}$
$q_k(s)$	Integer randomly generated from a uniform distribution within $[1, 35]$ for period $k \in \{-1, 0\}$ and repetition $s \in \{1, 2, \ldots, 500\}$
$I_0(s)$	$I_0(s) = \mu q_{-1}(s)\hat{\delta}_{-1,0}(s)$

τ. The horizontal arrows into the box $P(\tau)$ represent decisions made in the previous period $\tau - 1$, including the inventory level $I_{\tau-1}$, shortage level $g_{\tau-1}$, and the order decisions made in the past K periods, $\{q_{\tau-K}, q_{\tau-K+1}, \ldots, q_{\tau-1}\}$, where K is the upper bound on the uncertain lead time. These previously made decisions are used as parameters in $P(\tau)$.

3.4.2 Simulation results

Simulation results demonstrate that the tri-level model on average has lower total cost than other three deterministic models for a wide range of combinations of holding and shortage costs. To conduct a sensitivity analysis, cost parameters are changed, and the sample probability

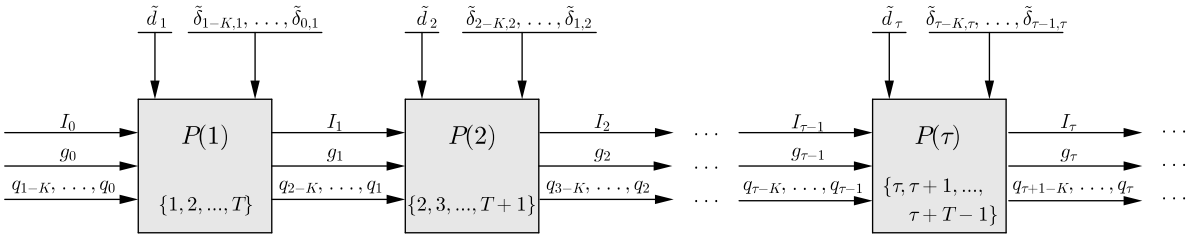


Figure 3.2: Planning horizons with length T in the simulation run

distribution of total cost for 500 repetitions and different combinations of cost parameters are shown in Figure 3.3. This Figure represents a combination of two cost parameters of h and p. Each row and column of graphs illustrate one shortage cost and one holding cost, respectively. The three deterministic models have different performance in response to cost parameters. As can be seen from Figure 3.3, the performance of Model 1 is improved when holding cost is high and shortage cost is low. In contrast, Model 3 works capably when holding cost is low and shortage cost is high; and Model 2 performs in between. However, the tri-level model works adaptively in response to changes in cost parameters and outperforms other models. Another point to consider, the result of the tri-level model and Model 3 are the same when inventory holding cost is very low; that is, $h = 1$. As shown in Figure 3.3, the proposed model has the lowest average cost, lowest worst solution, and lowest variance among all models in most combinations of cost parameters.

We show the results of the simulation in another perspective to illustrate how much the total cost of the tri-level model is better or worse than other models. The relative performance of the tri-level model compared to three deterministic models is evaluated by the ratio R = 100.(Det - Tri)/Det, where Det is the total cost of a deterministic model and Tri is the total cost of the tri-level model. The expectations of total costs are estimated on the repetition size of 500. We plot and show the results regarding the performance ratio of total cost in Figure 3.4. When the performance ratio is positive, the tri-level model works better than the compared model; thus, a higher percentage means a higher relative performance of the tri-level model. It is positive in all cases except the case with $h = 15$ and $p = 4$. When holding cost is very high and shortage cost is very low, Model 1 has a total cost lower than the tri-level model by an 11% average. Model 1 functions more effectively by increasing h or decreasing p

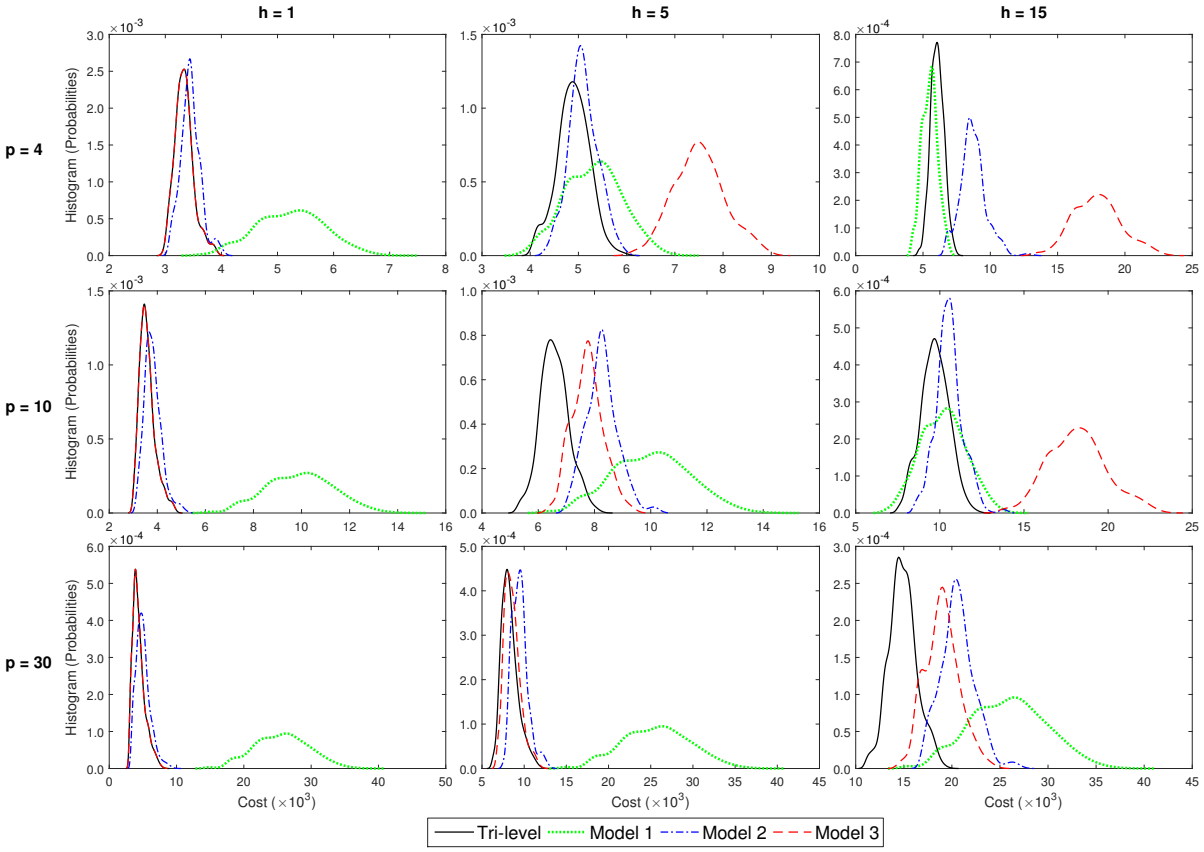


Figure 3.3: The sample probability distribution of the total costs for tri-level and three deterministic models with different holding and shortage costs

because if h is high or p is low, it is better to have a lower inventory level and possibly more shortages. Conversely, Model 3 works more poorly when h is increased and p is decreased. It tends to have a higher inventory level by forecasting future demands and lead times as large as possible; thus, the total cost of this model is raised by increasing h/p. In summary, the average performance ratios of the total cost over 500 repetitions and all nine cost parameter combinations for the proposed model with respect to Model 1 to 3 are 36%, 14%, and 21%, respectively. In terms of standard deviation of the total cost, the tri-level model averagely has a lower standard deviation compared to Models 1 to 3 by 58%, 10%, and 21%, respectively.

Furthermore, the tri-level model reacts adaptively to variation in cost parameters in terms of demand satisfaction rate. The percentage of customer orders satisfied immediately from stock at hand is called fill-rate. In general, it is improved by increasing shortage cost and decreasing holding cost. The average fill-rate of 500 repetitions for all models are shown in

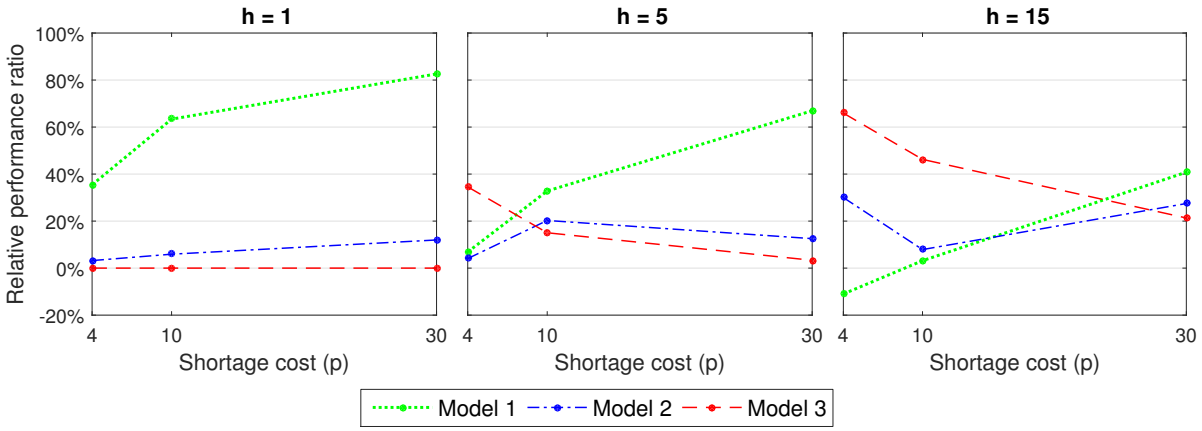


Figure 3.4: Impacts of holding and shortage costs on the relative performance ratio of the total cost

Figure 3.5. Fill-rate of Models 1 and 3 in all combinations of h and p are always equal to 69% and 97%, respectively. When holding cost equals 1, fill-rates of the tri-level model and Model 3 are the same for all different shortage costs. However, when holding cost is increased, the tri-level model responds to variation in shortage cost. The fill-rate of the tri-level model decreases significantly when holding cost is high and shortage cost is low.

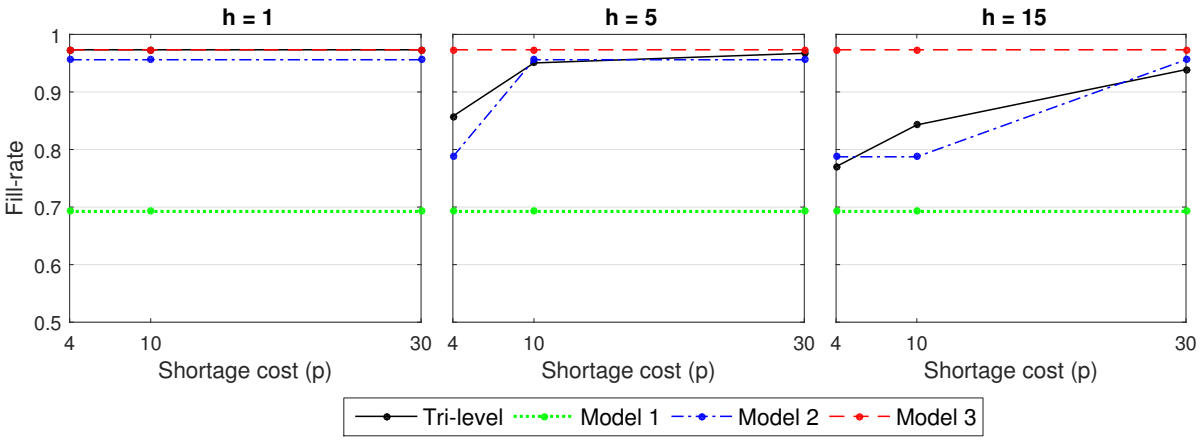


Figure 3.5: Impacts of holding and shortage costs on the fill-rate

To explain how our model outperforms the three deterministic models, we broke down the total cost into three parts for a randomly selected example. Results of the selected example for all nine cases and four models are summarized and shown in Table 3.4 and Figure 3.6.

The table reports the average amount of orders placed, average inventory levels, and average shortages in each period. The average orders for all models are almost the same, but the average inventory levels and shortages are different; the difference mostly comes from the time of orders. The results of Model 1 indicate that it tends to have zero inventory level and the maximum shortages among other models in all nine cases. When holding cost is 1 or 5 for all three shortage costs (the first six cases), the average inventory levels of the tri-level model are less than Model 2; thus, we may expect more shortage for the tri-level model. However, the average shortages of the tri-level model are also less than Model 2. It means that the tri-level model makes the right amount of orders at the right time. Complementary to this, when holding cost is high, $h = 15$, the tri-level model suggests having shortages more than Model 2 and 3 to manage the balance between holding and shortage costs. Figure 3.6 presents the average order cost, holding cost and shortage cost per period for four models and nine cases of this particular example through a stacked bar chart to compare the total cost of each model as well. Each bar in a group represents the costs of the tri-level model and Models 1 to 3 from left to right. The order cost for all models and all cases are approximately the same. The largest part of the total cost for Model 1 and 3 belongs to the shortage and inventory costs, respectively. Consider the case in Table 3.4 when $h = 15$ and $p = 4$. The tri-level model has a higher average shortage than Model 2, but it has a lower inventory level. As shown in Figure 3.6 the average shortage cost of the tri-level model is slightly more than Model 2, but the average inventory cost is considerably lower than Model 2. Therefore, the average total cost per period is lower than Model 2.

3.5 Conclusions

In this study, we propose a new approach to address uncertainty in a manufacturing facility which orders new items to satisfy demand. The demand and lead time are uncertain parameters, and shortages are fully backlogged. The objective is to make ordering decisions to minimize the total cost. This paper makes three contributions to the literature. First, we explicitly take into account two sources of uncertainty from both demand and lead time. Most previously proposed models focused on one of these two, but are still subject to significant uncertainty

Table 3.4: Average order, inventory level, and shortage in each period for one example and all models

	h	1			5			15		
	p	4	10	30	4	10	30	4	10	30
Average orders	Tri-level	34.83	34.83	34.83	33.73	34.40	34.67	33.27	33.67	34.27
	Model 1	32.77	32.77	32.77	32.77	32.77	32.77	32.77	32.77	32.77
	Model 2	34.47	34.47	34.47	33.30	34.47	34.47	33.30	33.30	34.47
	Model 3	34.83	34.83	34.83	34.83	34.83	34.83	34.83	34.83	34.83
Average inventory level	Tri-level	37.20	37.20	37.20	11.10	25.63	32.67	4.23	10.03	22.40
	Model 1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Model 2	39.17	39.17	39.17	11.87	39.17	39.17	11.87	11.87	39.17
	Model 3	37.20	37.20	37.20	37.20	37.20	37.20	37.20	37.20	37.20
Average shortage	Tri-level	0.00	0.00	0.00	4.40	0.60	0.03	11.37	5.47	0.90
	Model 1	22.73	22.73	22.73	22.73	22.73	22.73	22.73	22.73	22.73
	Model 2	0.63	0.63	0.63	7.17	0.63	0.63	7.17	7.17	0.63
	Model 3	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

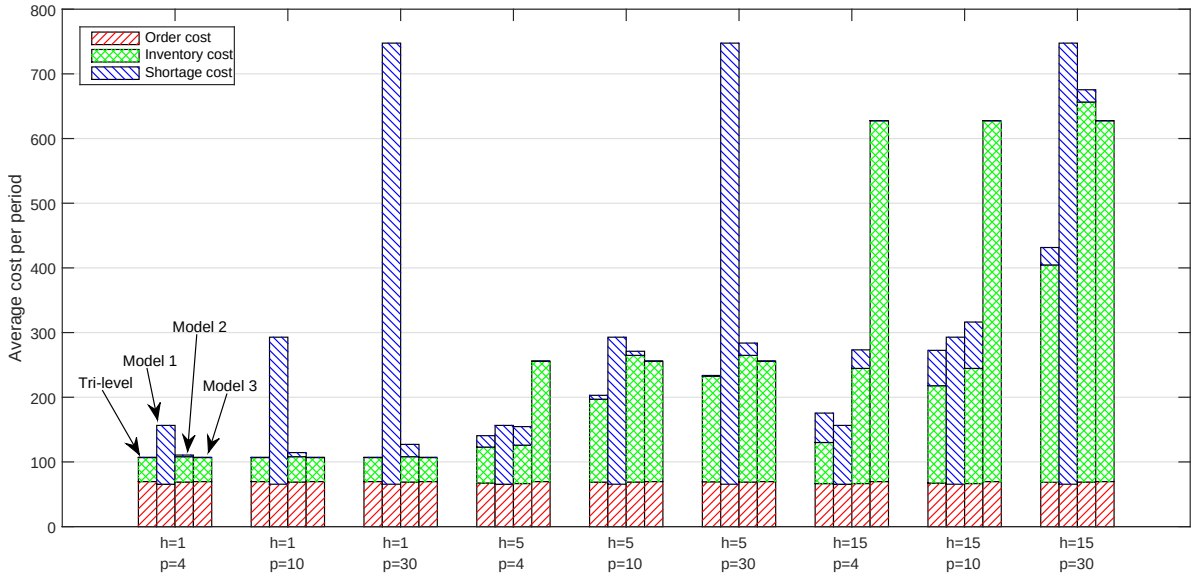


Figure 3.6: The average costs per period of four models and nine cases for one example

from the other source as well as the interactions of the two. Second, we propose a two-stage tri-level optimization model for the inventory control problem, which is a compromise between accurate representation of the multi-stage decision-making under uncertainty nature of the problem and computational tractability. Third, we design an exact algorithm for the tri-level optimization model, which deploys a Benders decomposition framework to efficiently search for the worst-case scenario without enumerating the enormous scenario space.

The results suggest that the tri-level optimization model works more adaptively in response to a wide range of cost parameters. The performances of Models 1 and 3 are highly sensitive to the cost parameters, and Model 2 is almost always in between. In contrast, the tri-level optimization model automatically adjusts its optimal ordering strategies according to the cost parameters and yields the lowest (or close to lowest) total cost for all parameter settings.

This study is subject to several limitations which suggest future research directions. For example, the proposed model assumes a single product made from a single part. Relaxing this assumption would require a more complicated model that reflects the uncertainty and interdependency of multiple parts on the demand and supply sides. In addition, we can include fixed and variable transportation costs in the model, where the decision maker has the option to ship certain parts or products together as a batch to save transportation cost.

BIBLIOGRAPHY

Axsäter, S. (2003). Approximate optimization of a two-level distribution inventory system. *International Journal of Production Economics*, 81:545–553.

Beaudin, M. and Zareipour, H. (2015). Home energy management systems: A review of modelling and complexity. *Renewable and Sustainable Energy Reviews*, 45:318–335.

Bertsimas, D. and Thiele, A. (2006). Robust and data-driven optimization: Modern decision-making under uncertainty. *INFORMS Tutorials in Operations Research*, pages 95–122.

Davis, T. (1993). Effective supply chain management. *Sloan Management Review*, 34:35–46.

Defourny, B., Ernst, D., and Wehenkel, L. (2011). Multistage stochastic programming: A scenario tree based approach. *Decision Theory Models for Applications in Artificial Intelligence: Concepts and Solutions*, pages 97–143.

Dey, J. K., Mondal, S. K., and Maiti, M. (2008). Two storage inventory problem with dynamic demand and interval valued lead-time over finite time horizon under inflation and time-value of money. *European Journal of Operational Research*, 185(1):170–194.

Ho, C. F., Chi, Y. P., and Tai, Y. M. (2005). A structural approach to measuring uncertainty in supply chains. *International Journal of Electronic Commerce*, 9(3):91–114.

Hoque, M. (2013). A manufacturer–buyer integrated inventory model with stochastic lead times for delivering equal-and/or unequal-sized batches of a lot. *Computers & Operations Research*, 40(11):2740–2751.

Li, L. and Schulze, L. (2011). Uncertainty in logistics network design: a review. In *Proceedings of the International MultiConference of Engineers and Computer Scientists*, volume 2.

Maiti, A., Maiti, M., and Maiti, M. (2009). Inventory model with stochastic lead-time and price dependent demand incorporating advance payment. *Applied Mathematical Modelling*, 33(5):2433–2443.

Muriana, C. (2016). An eoq model for perishable products with fixed shelf life under stochastic demand conditions. *European Journal of Operational Research*, 255(2):388–396.

Pan, F. and Nagi, R. (2010). Robust supply chain design under uncertain demand in agile manufacturing. *Computers & Operations Research*, 37(4):668–683.

Routroy, S. and Kodali, R. (2005). Differential evolution algorithm for supply chain inventory planning. *Journal of Manufacturing Technology Management*, 16(1):7–17.

Sahin, F., Narayanan, A., and Robinson, E. P. (2013). Rolling horizon planning in supply chains: review, implications and directions for future research. *International Journal of Production Research*, 51(18):5413–5436.

Sajadieh, M. S., Jokar, M. R. A., and Modarres, M. (2009). Developing a coordinated vendor–buyer model in two-stage supply chains with stochastic lead-times. *Computers & Operations Research*, 36(8):2484–2489.

Seifbarghy, M. and Jokar, M. R. A. (2006). Cost evaluation of a two-echelon inventory system with lost sales and approximately poisson demand. *International Journal of Production Economics*, 102(2):244–254.

Unlu, Y. and Rossetti, M. D. (2009). Evaluating the lead time demand distribution for (r, q) policies under intermittent demand. In *IIE Annual Conference. Proceedings*, pages 1670–1675. Institute of Industrial Engineers-Publisher.

Wang, J. and Shu, Y. F. (2005). Fuzzy decision modeling for supply chain management. *Fuzzy Sets and Systems*, 150(1):107–127.

Wang, Q. (2013). A periodic-review inventory control policy for a two-level supply chain with multiple retailers and stochastic demand. *European Journal of Operational Research*, 230(1):53–62.

Xu, P. and Wang, L. (2014). An exact algorithm for the bilevel mixed integer linear programming problem under three simplifying assumptions. *Computers & Operations Research*, 41:309–318.

CHAPTER 4. A NEW BRANCH AND BOUND ALGORITHM FOR THE BILEVEL LINEAR PROGRAMMING PROBLEM

A paper prepared to submit to IIE Transactions

Mohammad Rahdar and Lizhi Wang

Abstract

This paper presents a new branch and bound algorithm to solve a bilevel linear programming problem, which is proven to be NP-hard. First, we replace the follower problem with its "Karush-Kuhn-Tucker" conditions to reformulate the two-level mathematical model as a single-level model with linear complementarity constraints (LPCC). Then, we solve the LPCC problem in a branch and bound scheme to satisfy the complementary slackness conditions. The proposed algorithm is examined by solving 100 randomly generated instances with different sizes and compared to the original branch and bound algorithm. The results indicate that the new algorithm is more efficient than the original one.

4.1 Introduction

Bilevel Linear Programming (BLP) problems are a special case of optimization problems with two decision makers, as refer to the leader (upper-level) and the follower (lower-level); thus, there are two types of variables referred to two levels. The optimization problem of the follower is enclosed within the constraints of the leader problem. Therefore, the solution of the upper-level problem is feasible only if this solution is optimal to the lower-level problem. A generic formulation of bi-level programming problem is as follows.

$$\max_{x,y} \quad c^\top x + d_1^\top y \tag{4.1}$$

$$\text{s.t.} \quad A_1 x + B_1 y \leq b_1 \tag{4.2}$$

$$y \in \arg\max_{\tilde{y}} \{d_2^\top \tilde{y} : A_2 x + B_2 \tilde{y} \leq b_2\} \tag{4.3}$$

Where $A_1 \in \mathbb{R}^{m_1 \times n_1}$, $B_1 \in \mathbb{R}^{m_1 \times n_2}$, $A_2 \in \mathbb{R}^{m_2 \times n_1}$, $B_2 \in \mathbb{R}^{m_2 \times n_2}$, $b_1 \in \mathbb{R}^{m_1 \times 1}$, $b_2 \in \mathbb{R}^{m_2 \times 1}$, $c \in \mathbb{R}^{n_1 \times 1}$, $d_1 \in \mathbb{R}^{n_2 \times 1}$, and $d_2 \in \mathbb{R}^{n_2 \times 1}$. Since the leader has full knowledge of the follower's problem, it attempts to maximize its objective function by selecting a strategy that foresees the reactions of the follower. The upper-level decision variable is x and the lower-level variable is y. First, the leader anticipates the response of the follower and decides on variable x to optimize its objective function; then, the follower optimizes its objective by deciding on variable y. It is the optimistic bi-level linear programming problem because if the follower has multiple optimal solutions, the leader can choose the one that optimizes its own objective function. Bard and Moore (1990) stated that bilevel programming has two main assumptions: first, both players have access to the full information; second, cooperation is not allowed. This makes bilevel optimization problems difficult to solve. Many researchers have studied the properties of the BLP problems. Bard (1991), Ben-Ayed and Blair (1990), Hansen et al. (1992), and Vicente et al. (1994) proved that the BLP problem is NP-hard and discussed the difficulties of developing efficient algorithms to solve it. However, it has many important application areas, such as economics, transportation, and business. Due to the lack of efficient algorithms for tackling medium and large scale bi-level programming problems, we have been motivated to work on developing a new algorithm or improving current methods to solve these types of problems more efficiently.

4.1.1 Definitions

In this section, we provide some definitions, which were given by Bard (2013).

(a) Constraint region of the BLP problem:

$$S = \{(x, y) : A_1 x + B_1 y \leq b_1, A_2 x + B_2 y \leq b_2\} \tag{4.4}$$

Set S represents the feasible region of both upper and lower-level constraints.

(b) Feasible set for the follower for each fixed x:

$$S(x) = \{y : B_2 y \leq b_2 - A_2 x\} \tag{4.5}$$

Feasible set of the follower is affected by the leader's choice of x.

(c) Projection of S onto the leader's decision space:

$$X = \{x : \exists y \text{ such that } A_1 x + B_1 y \leq b_1, A_2 x + B_2 y \leq b_2\} \tag{4.6}$$

The range of x that leader can choose from.

(d) Follower's rational reaction set for $x \in X$:

$$P(x) = \{y : y \in \arg\max_{\tilde{y}}[f(x, \tilde{y}) : \tilde{y} \in S(x)]\} \tag{4.7}$$

The follower reacts based on the leader's action, and select y from its feasible set $S(x)$.

(e) Inducible region:

$$\text{IR} = \{(x, y) : (x, y) \in S, y \in P(x)\} \tag{4.8}$$

Inducible region is the feasible set of the BLP problem. Therefore, the BLP problem (4.1)-(4.3) is equivalent to (4.9).

$$\max_{x,y}\{c^\top x + d_1^\top y : (x, y) \in \text{IR}\} \tag{4.9}$$

4.1.2 Categorizing BLP into seven different cases

We provide seven examples, which are easy to solve manually, to represent seven types of the bilevel programming problems.

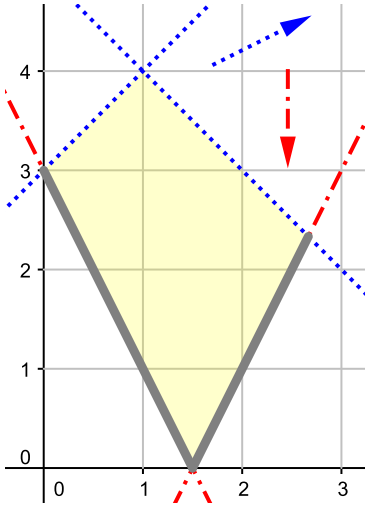
- The dotted and the dash-dot lines show the constraints of the upper-level and the lower-level problem, respectively.
- The dotted and the dash-dot vectors indicate the improvement direction of objective functions of the upper and lower-level problems, respectively.

- The shaded area represents the feasible region of the relaxation problem, Equation (4.4).
- The thick solid lines indicate the optimal solution of the lower-level problem for given variable x; that is, the follower rational reaction set $P(x)$, Equation (4.7).

In Examples 1 and 2, both the relaxation and lower-level problems have an optimal solution, but the BLP problem may have an optimal solution or be infeasible.

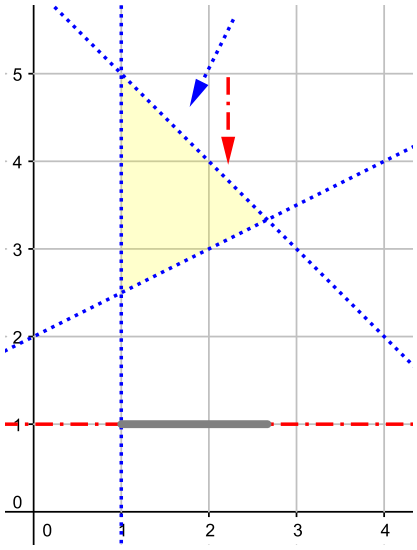
Example 1. The following BLP problem has an optimal solution ($x^* = 2.67, y^* = 2.33$). Since the constraint region of the BLP problem (shaded area) is compact, the relaxation problem has an optimal solution. In addition, for any value of x in the range of 0 and 2.67, the lower-level problem has an optimal solution, which is shown by a thick solid line.

$$
\begin{aligned}
\max \quad & 2x + y \\
\text{s.t.} \quad & -x + y \leq 3 \\
& x + y \leq 5 \\
& \max \quad -y \\
& \text{s.t.} \quad -2x - y \leq -3 \\
& \qquad 2x - y \leq 3
\end{aligned}
$$



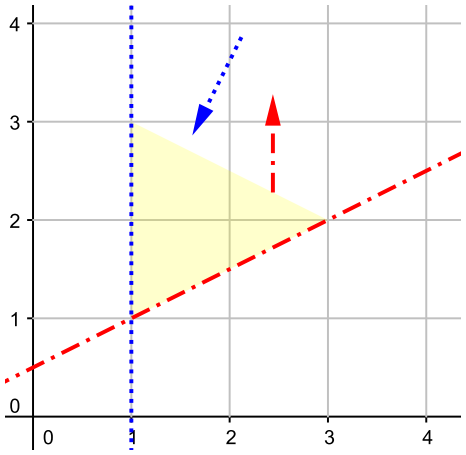
Example 2. The following BLP is infeasible, although, both the relaxation and lower-level problems have an optimal solution. The optimal solution of the lower-level problem for any given x within the range of 1 and 2.67 is outside the constraint region of the BLP problem; thus, the BLP does not have any feasible solution.

$$
\begin{aligned}
\max \quad & -x - 2y \\
\text{s.t.} \quad & -x \leq -1 \\
& x - 2y \leq -4 \\
& x + y \leq 6 \\
& \max \ -y \\
& \text{s.t.} \quad -y \leq -1
\end{aligned}
$$



Example 3. The following BLP is infeasible because no matter what value x takes, the lower-level is unbounded.

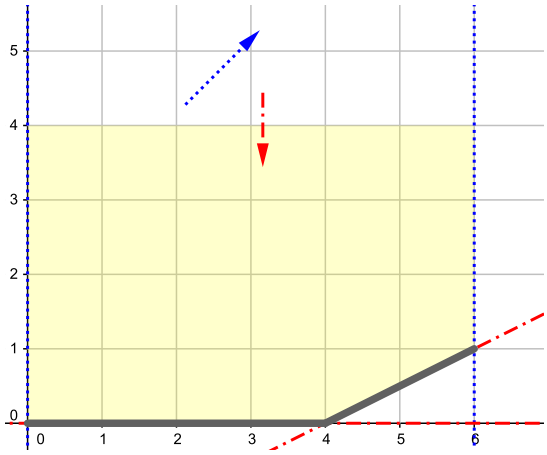
$$
\begin{aligned}
\max \quad & -x - 2y \\
\text{s.t.} \quad & -x \leq -1 \\
& \max \ y \\
& \text{s.t.} \quad x - 2y \leq -1
\end{aligned}
$$



In Examples 4 and 5 the relaxation problem is unbounded, and the lower-level problem has an optimal solution, but the BLP problem may have an optimal solution or be unbounded.

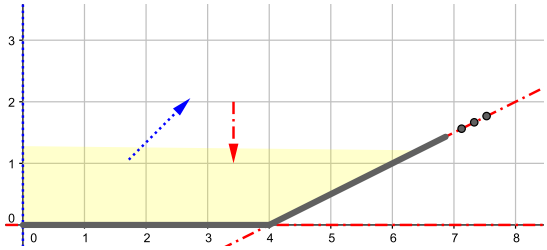
Example 4. The following BLP problem has an optimal solution $(x^* = 6, y^* = 1)$.

$$
\begin{aligned}
\max \quad & x + y \\
\text{s.t.} \quad & x \leq 6 \\
& -x \leq 0 \\
& \max \quad -y \\
& \text{s.t.} \quad x - 2y \leq 4 \\
& \qquad -y \leq 0
\end{aligned}
$$



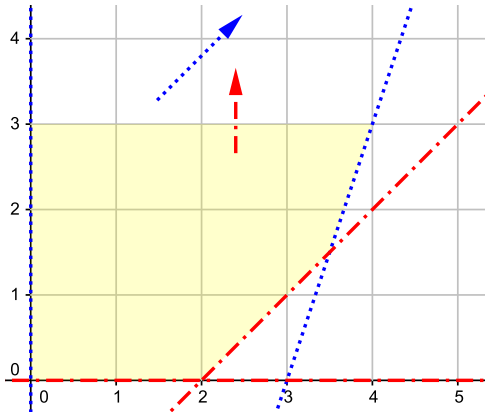
Example 5. Similar to Example 4, the relaxation problem is unbounded, and the lower-level problem has an optimal solution. However, the BLP problem is unbounded considering the unboundedness of inducible region and the improvement direction of upper-level objective function.

$$
\begin{aligned}
\max \quad & x + y \\
\text{s.t.} \quad & -x \leq 0 \\
& \max \quad -y \\
& \text{s.t.} \quad x - 2y \leq 4 \\
& \qquad -y \leq 0
\end{aligned}
$$



Example 6. The following BLP problem is infeasible because similar to Example 3 the follower problem is unbounded.

$$
\begin{aligned}
\max \quad & x + y \\
\text{s.t.} \quad & 3x - y \leq 9 \\
& -x \leq 0 \\
& \max \quad y \\
& \text{s.t.} \quad x - y \leq 2 \\
& \qquad -y \leq 0
\end{aligned}
$$



Example 7. The following BLP problem is infeasible because the relaxation problem is infeasible.

$$
\begin{aligned}
\max \quad & x + y \\
\text{s.t.} \quad & x - y \leq -3 \\
& -x \leq 0 \\
& \max \quad y \\
& \text{s.t.} \quad y \leq 2
\end{aligned}
$$

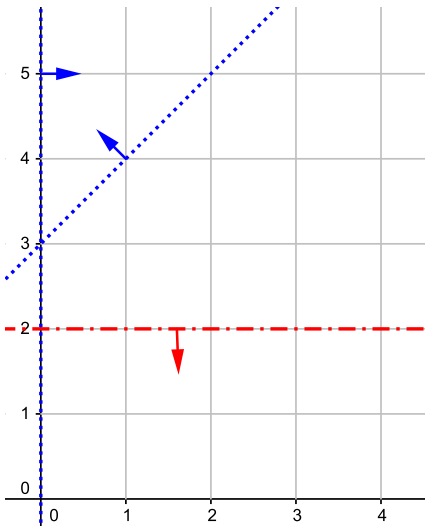


Table 4.1 represents the summary of seven possible types of BLP problem.

Table 4.1: Seven possible types of BLP problem

Example	Relaxation problem	Lower-level problem	BLP
1	Optimal	Optimal	Optimal
2	Optimal	Optimal	Infeasible
3	Optimal	Unbounded	Infeasible
4	Unbounded	Optimal	Optimal
5	Unbounded	Optimal	Unbounded
6	Unbounded	Unbounded	Infeasible
7	Infeasible	Any	Infeasible

4.2 Literature review

Many researchers have studied bilevel optimization and developed algorithms to solve this type of problems. However, most of these algorithms are not applicable due to computational limitations and simplifying assumptions, which they need. As stated in Bard (2013), there are generally three different approaches to solving bilevel linear programming problems. The first type of approaches use some form of vertex enumeration, the second one applies penalty approach, and finally, the third type of methods involves the "Karush-Kuhn-Tucker" (KKT) conditions to convert the bilevel programming problem to a single level problem [Bard (2013)].

Several studies are based on vertex enumeration method to explore bases of the constraint region. Bialas and Karwan (1982) introduced the Kth-best method in which the algorithm searches for an optimal solution among extreme points of the constraint region S. Another class of algorithms is based on some form of penalty approaches. Aiyoshi and Shimizu (1984) solved the BLP problem by first converting the follower problem to an unconstrained mathematical program; then, replacing the penalized problem of the follower by its stationary condition. White and Anandalingam (1993) used a penalty function for the duality gap of the follower's problem in the leader's objective function. When the duality gap for a value of x becomes zero, the solution y is optimal to the follower problem. Thus, it would be in the rational reaction set.

One direct approach to solving a bilevel linear programming model is reformulating it as a linear program with linear complementarity constraints (LPCC) by applying KKT conditions. Bard and Moore (1990) designed a branch and bound algorithm to find a solution which satisfies the complementarity term. Fortuny-Amat and McCarl (1981) converted the LPCC problem into a mixed-integer linear program by adding a binary variable, a sufficiently large constant M, and two constraints to the model. Furthermore, Hu et al. (2008) introduced a big-M-free algorithm to solve LPCC by applying Benders decomposition method.

In addition to global optimization techniques, many researchers have developed heuristic and artificial intelligence-based algorithms to solve bilevel linear programming problems. Mathieu et al. (1994), Yin (2000), Oduguwa and Roy (2002), Hejazi et al. (2002), Wang et al.

(2005) Calvete et al. (2008), and Osman et al. (2009) proposed solving techniques based on genetic algorithm. Other algorithms which have been developed to solve BLP problems are simulated annealing [Sahin and Ciric (1998)], tabu search [Gendreau et al. (1996) and Rajesh et al. (2003)], and particle swarm optimization [Kuo and Huang (2009)] to solve BLP problems. Moreover, a one-dimensional grid search algorithm was developed by Bard (1983); they showed that their algorithm is convergent under fairly general conditions.

The most commonly used algorithms in the third approach, stated by Bard (2013), are Branch and Bound, Big-M, and Benders algorithms. In this section, we explain and summarize these algorithms in more detail.

4.2.1 Branch and Bound

Bard and Moore (1990) proposed a branch and bound algorithm to solve BLP problems. They suggested to convert the hierarchical problem to a single level standard mathematical program with complementarity constraints by applying "Karush-Kuhn-Tucker" conditions, and then solve this problem by branch and bound method. The equivalent formulation of the BLP problem is achieved by replacing the lower-level problem (4.3) with its KKT conditions as follows:

$$\max_{x,y,\lambda} \quad \zeta = c^\top x + d_1^\top y \tag{4.10}$$

$$\text{s. t.} \quad A_1 x + B_1 y \leq b_1 \tag{4.11}$$

$$0 \leq b_2 - A_2 x - B_2 y \perp \lambda \geq 0 \tag{4.12}$$

$$B_2^\top \lambda = d_2 \tag{4.13}$$

where $(b_2 - A_2x - B_2y \perp \lambda)$ means that the two vectors are orthogonal; in other words, $(b_2 - A_2x - B_2y)^\top \lambda = 0$. Since there are m_2 complementarity constraints, there are 2^{m_2} subsets, or equivalently 2^{m_2} linear programs. Therefore, this formulation is equivalent to solve a large number of linear programs. The optimal solution is the best among the solution of 2^{m_2} linear programs. Branch and bound method looks for the optimal solution by branching on complementarity constraint $i, \forall i = 1, 2, ..., m_2$. That is, one branch makes $(b_2 - A_2x - B_2y)_i = 0$ and another one makes $\lambda_i = 0$.

The branch and bound algorithm cannot solve BLP problems when the relaxation problem or lower-level problem is unbounded. In other words, it cannot solve problem types 3 to 6 in Table 4.1. Bard and Moore (1990) assumed that the constraint region (4.4) is nonempty and compact.

For given parameters u and v, we define $\mathcal{R}(u, v)$ as the following parametric LP relaxation:

$$\max_{x,y,\lambda} \quad \zeta = c^\top x + d_1^\top y \tag{4.14}$$

$$\text{s.t.} \quad A_1 x + B_1 y \leq b_1 \tag{4.15}$$

$$0 \leq b_2 - A_2 x - B_2 y \leq u \tag{4.16}$$

$$0 \leq \lambda \leq v \tag{4.17}$$

$$B_2^\top \lambda = d_2 \tag{4.18}$$

Algorithm 4.1 Branch and Bound

function BLP-BB(A_1, A_2, B_1, B_2, b_1, b_2, c, d_1, d_2)

Step 0: Create node 1, which is characterized by $(U^1 = \infty^{m_2 \times 1}, V^1 = \infty^{m_2 \times 1}, Z^1 = \infty)$. Initialize $x^* = \emptyset$, $y^* = \emptyset$, $\zeta^* = -\infty$, and $N = 1$. Go to Step 1.

Step 1: For all $j \in \{1, ..., N\}$ such that $Z^j \leq \zeta^*$, discard node j. Update N as the number of remaining nodes.

  if $N = 0$ **then**

    if $x^* \neq \emptyset$ **then**

      1(a) return (x^*, y^*, ζ^*) is an optimal solution to the BLP (4.1)-(4.3).

    else

      1(b) return BLP (4.1)-(4.3) is infeasible.

    end if

  else

    1(c) select a node k from $\{1, ..., N\}$, set $\hat{u} = U^k$ and $\hat{v} = V^k$, discard node k, reorder the remaining nodes from 1 to $N - 1$, reduce N by 1, and go to Step 2.

  end if

Step 2: Solve $\mathcal{R}(\hat{u}, \hat{v})$.

  if $\mathcal{R}(\hat{u}, \hat{v})$ is infeasible **then**

    2(a) go to Step 1.

  else

    Let $(x^{\mathrm{R}}, y^{\mathrm{R}}, \lambda^{\mathrm{R}})$ denote an optimal solution and ζ^{R} the optimal objective value of $\mathcal{R}(\hat{u}, \hat{v})$.

Algorithm 4.1 Branch and Bound (continued)

if $\zeta^{\mathrm{R}} \leq \zeta^*$ **then**
  2(b) go to Step 1.
else if $(b_2 - A_2x^{\mathrm{R}} - B_2y^{\mathrm{R}}) \perp \lambda^{\mathrm{R}}$ **then**
  2(c) update $x^* = x^{\mathrm{R}}$, $y^* = y^{\mathrm{R}}$, and $\zeta^* = \zeta^{\mathrm{R}}$ and go to Step 1.
else
  2(d) create two new nodes, characterized by $(U^{N+1} = \hat{u}, V^{N+1} = \hat{v}, Z^{N+1} = \zeta^{\mathrm{R}})$ and $(U^{N+2} = \hat{u}, V^{N+2} = \hat{v}, Z^{N+2} = \zeta^{\mathrm{R}})$. Select $i \in \{1, ..., m_2\}$ such that $(b_2 - A_2x^{\mathrm{R}} - B_2y^{\mathrm{R}})_i > 0$ and $\lambda_i^{\mathrm{R}} > 0$. Change $U_i^{N+1} = 0$ and $V_i^{N+2} = 0$, increase N by 2, and go to Step 1.
end if
end if
end function

4.2.2 Big-M method

Similar to the previous method, we convert the two-level problem to a single-level model by applying "Karush-Kuhn-Tucker" conditions to obtain the formulation (4.10)-(4.13). Fortuny-Amat and McCarl (1981) addressed the complementarity constraints by adding a binary variable $z \in \mathbb{R}^{m_2 \times 1}$ and a large enough positive constant to the model. The alternative formulation of (4.10)-(4.13) is shown in (4.19)-(4.25).

$$\max_{x,y,z,\lambda} \quad c^\top x + d_1^\top y \tag{4.19}$$

$$\text{s.t.} \quad A_1x + B_1y \leq b_1 \tag{4.20}$$

$$A_2x + B_2y \leq b_2 \tag{4.21}$$

$$-A_2x - B_2y \leq -b_2 + Mz \tag{4.22}$$

$$\lambda \leq M(1 - z) \tag{4.23}$$

$$B_2^\top \lambda = d_2 \tag{4.24}$$

$$\lambda \geq 0 \tag{4.25}$$

Problem (4.19)-(4.25) is a mixed-integer programming problem, which can be solved by current solvers and algorithms. However, estimating M is difficult. It is hard to know how big is enough; if the constant M is too small, it can eliminate the optimal solution. If it is too large, it may cause computational errors. Therefore, using this algorithm has practical issues and difficulties.

4.2.3 Benders algorithm

Hu et al. (2008) developed a big-M-free algorithm to solve a linear program with linear complementarity constraints. They proposed a logical Benders scheme to find the global solution. This algorithm was the first one, which could solve all possible types of BLP problems, but it is very slow even for small problems. We modified their algorithm to solve the BLP problems, but before introducing it, we need to define three problems.

$$(\mathbf{M0}) \min_{z} \quad 0 \tag{4.26}$$

$$\text{s. t.} \quad Ez \geq h \tag{4.27}$$

$$z \in \mathbb{B}^{m_2 \times 1} \tag{4.28}$$

$$(\mathbf{SE0}) \min_{\Delta\lambda,\Delta u,\Delta w,\Delta v,\Delta\beta} \quad b_1^\top \Delta\lambda + b_2^\top(\Delta u - \Delta w) + d_2^\top \Delta\beta \tag{4.29}$$

$$\text{s. t.} \quad A_1^\top \Delta\lambda + A_2^\top(\Delta u - \Delta w) = 0 \tag{4.30}$$

$$B_1^\top \Delta\lambda + B_2^\top(\Delta u - \Delta w) = 0 \tag{4.31}$$

$$\Delta v + B_2 \Delta\beta \geq 0 \tag{4.32}$$

$$(z^{\mathrm{M}})^\top \Delta w + (1 - z^{\mathrm{M}})^\top \Delta v = 0 \tag{4.33}$$

$$0 \leq \Delta\lambda, \Delta u, \Delta w, \Delta v \leq 1; -1 \leq \Delta\beta \leq 1 \tag{4.34}$$

$$(\mathbf{SE}) \min_{\lambda,u,w,v,\beta} \quad b_1^\top \lambda + b_2^\top(u - w) + d_2^\top \beta \tag{4.35}$$

$$\text{s. t.} \quad A_1^\top \lambda + A_2^\top(u - w) = c \qquad (x) \tag{4.36}$$

$$B_1^\top \lambda + B_2^\top(u - w) = d_1 \qquad (y) \tag{4.37}$$

$$v + B_2 \beta \geq 0 \qquad (\lambda) \tag{4.38}$$

$$(z^{\mathrm{M}})^\top w + (1 - z^{\mathrm{M}})^\top v = 0 \tag{4.39}$$

$$\lambda, u, w, v \geq 0 \tag{4.40}$$

Algorithm 4.2 Benders Algorithm

function BLP-BENDERS(A_1, A_2, B_1, B_2, b_1, b_2, c, d_1, d_2)

Step 0: Initialize an empty matrix $E \in \mathbb{R}^{0\times m_2}$, an empty vector $h \in \mathbb{R}^{0\times 1}$, $x^* = \emptyset$, $y^* = \emptyset$, and $\zeta^* = -\infty$.

Step 1: Solve (**M0**).

  if is infeasible **then**
    if $\zeta^* = -\infty$ **then**
      1(a) return BLP is infeasible.
    else
      1(b) return Solution (x^*, y^*, ζ^*) is optimal to BLP .
    end if
  else
    1(c) let z^{M} denote a feasible solution to and go to Step 2.
  end if

Step 2: Solve (**SE0**) and let $(\Delta\lambda^0, \Delta u^0, \Delta w^0, \Delta v^0, \Delta\beta^0)$ be an optimal solution.

  if $b_1^\top \Delta\lambda^0 + b_2^\top(\Delta u^0 - \Delta w^0) + d_2^\top \Delta\beta^0 < 0$ **then**
    2(a) Update $E \leftarrow \begin{bmatrix} E \\ (\Delta w^0 - \Delta v^0)^\top \end{bmatrix}$, $h \leftarrow \begin{bmatrix} h \\ \epsilon(\Delta w^0, \Delta v^0) - \sum_{i=1}^{m_2} \Delta v_i^0 \end{bmatrix}$, where $\epsilon(\Delta w^0, \Delta v^0)$ is the smallest positive element in Δw^0 and Δv^0. Go to Step 1.
  else
    Solve (**SE**).
    if is infeasible **then**
      2(b) return : BLP (4.1)-(4.3) is unbounded.
    else
      2(c) Let $(\lambda^{\mathrm{E}}, u^{\mathrm{E}}, w^{\mathrm{E}}, v^{\mathrm{E}}, \beta^{\mathrm{E}})$ be an optimal solution to and $(x^{\mathrm{E}}, y^{\mathrm{E}}, \lambda^{\mathrm{E}})$ its optimal dual solution.
      if $c^\top x^{\mathrm{E}} + d_1^\top y^{\mathrm{E}} > \zeta^*$ **then**
        Update $x^* = x^{\mathrm{E}}$, $y^* = y^{\mathrm{E}}$, and $\zeta^* = c^\top x^* + d_1^\top y^*$.
      end if
      Update $E \leftarrow \begin{bmatrix} E \\ (w^* - v^*)^\top \end{bmatrix}$, $h \leftarrow \begin{bmatrix} h \\ \epsilon(w^*, v^*) - \sum_{i=1}^{m_2} v_i^* \end{bmatrix}$, where $\epsilon(w^*, v^*)$ is the smallest positive element in w^* and v^*. Go to Step 1.
    end if
  end if

end function

4.3 The new branch and bound algorithm

As explained in Section 4.2.1, the formulation (4.10)-(4.13) is equivalent to the BLP problem (4.1)-(4.3). One approach to solve the mathematical program (4.10)-(4.13) is using the branch and bound method [Bard and Moore (1990)] by branching on the complementarity constraint (4.12). After relaxing the complementarity constraints and adding parameters u and v, we can define the following parametric relaxation problem $\mathcal{R}(u, v)$ as follows. Parameters u and v can be infinity or zero. $u_i = 0$ means that the ith constraint in (4.43) is binding. In the same way, if $v_i = 0$, then $\lambda_i = 0$.

$$\max_{x,y,\lambda} \quad \zeta = c^\top x + d_1^\top y \tag{4.41}$$

$$\text{s.t.} \quad A_1 x + B_1 y \leq b_1 \tag{4.42}$$

$$0 \leq b_2 - A_2 x - B_2 y \leq u \tag{4.43}$$

$$0 \leq \lambda \leq v \tag{4.44}$$

$$B_2^\top \lambda = d_2 \tag{4.45}$$

If the solution of the relaxation problem does not satisfy complementary slackness conditions, the selected node is branched into two new nodes to satisfy at least one complementarity constraint. One branch makes $u_i = 0$ and another one makes $v_i = 0$. If we update the parameter u in constraint (4.43) for the first child node, the solution of the parent node for λ is still valid. In the same way, when we update the parameter v in constraint (4.44) for the second child node, the solution of the parent node for (x, y) is still feasible. Therefore, we do not need to solve the whole relaxation problem $\mathcal{R}(u, v)$ in each node, and we can subdivide it into two smaller problems, such that the integration of them is equivalent to the problem $\mathcal{R}(u, v)$.

For the node with $u_i = 0$, we are looking for $(x^{\mathrm{R}}, y^{\mathrm{R}})$ by enforcing the ith constraint of the lower-level problem binding. We just need to solve the problem $\mathcal{P}(u)$, and use the solution of the parent node for the value of variable λ. Problem $\mathcal{P}(u)$ is as follows:

$$\max_{x,y} \quad c^\top x + d_1^\top y \tag{4.46}$$

$$\text{s.t.} \quad A_1 x + B_1 y \leq b_1 \tag{4.47}$$

$$0 \leq b_2 - A_2 x - B_2 y \leq u \tag{4.48}$$

Similarly, for the node with $v_i = 0$, we are looking for λ^{R} such that the ith dual variable becomes zero. The solution of the parent node for variable (x, y) is still feasible to this node. Thus, we only need to solve problem $\mathcal{D}(v)$ as follows:

$$\max_{\lambda} \quad 0 \tag{4.49}$$

$$\text{s. t.} \quad 0 \leq \lambda \leq v \tag{4.50}$$

$$B_2^{\top}\lambda = d_2 \tag{4.51}$$

The proposed algorithm does not solve the whole problem $\mathcal{R}(u, v)$ in each node; rather, it only solves the problem $\mathcal{P}(u)$ or the problem $\mathcal{D}(v)$. If the branch makes $u_i = 0$, it solves the problem $\mathcal{P}(u)$ to determine (x, y), and uses the solution of the parent node for determining the dual variable of the lower-level problem (λ). In the same way, if the branch makes $v_i = 0$, it solves the problem $\mathcal{D}(v)$ to find λ, and uses the solution of the parent node for determining (x, y). Therefore, a smaller problem is solved in each node, which can shorten the solving time considerably.

Similar to Bard and Moore (1990), we assume that the constraint region (4.4) is nonempty and compact. In addition, the follower can always respond to each decision of the leader. The new algorithm is as follows.

Algorithm 4.3 The New Branch and Bound Algorithm

function BLP-NewAlg(A_1, A_2, B_1, B_2, b_1, b_2, c, d_1, d_2)

  Step 1: solve $\mathcal{R}(\infty, \infty)$.

    if $\mathcal{R}(\infty, \infty)$ is infeasible **then**

      1.a return BLP (4.1)-(4.3) is infeasible.

    else

      Let $(x^{\mathrm{R}}, y^{\mathrm{R}}, \lambda^{\mathrm{R}})$ denote an optimal solution and ζ^{R} the optimal objective value of $\mathcal{R}(\infty, \infty)$.

      if $(b_2 - A_2x^{\mathrm{R}} - B_2y^{\mathrm{R}}) \perp \lambda^{\mathrm{R}}$ **then**

        1.b return $(x^{\mathrm{R}}, y^{\mathrm{R}}, \zeta^{\mathrm{R}})$ is an optimal solution to the BLP (4.1)-(4.3).

      else

        1.c Initialize $x^* = \emptyset$, $y^* = \emptyset$, $\zeta^* = -\infty$, $\hat{u} = \infty$, $\hat{v} = \infty$, and $N = 0$. Go to Step 2.

      end if

    end if

Algorithm 4.3 The new Branch and Bound algorithm (continued)

Step 2: Create two new nodes, characterized by
$(U^{N+1} = \hat{u}, V^{N+1} = \hat{v}, Z^{N+1} = \zeta^{\mathrm{R}}, X^{N+1} = \emptyset, Y^{N+1} = \emptyset, \Lambda^{N+1} = \lambda^{\mathrm{R}})$ and
$(U^{N+2} = \hat{u}, V^{N+2} = \hat{v}, Z^{N+2} = \zeta^{\mathrm{R}}, X^{N+2} = x^{\mathrm{R}}, Y^{N+2} = y^{\mathrm{R}}, \Lambda^{N+1} = \emptyset)$.
Select $i = \arg\max_{i \in \{1,\dots,m_2\}} \left(\lambda_i^{\mathrm{R}} \cdot (b_2 - A_2 x^{\mathrm{R}} - B_2 y^{\mathrm{R}})_i\right)$.
Change $U_i^{N+1} = 0$ and $V_i^{N+2} = 0$, increase N by 2, and go to Step 3.

Step 3: For all $j \in \{1, ..., N\}$ such that $Z^j \leq \zeta^*$, discard node j. Update N as the number of remaining nodes.

if $N = 0$ **then**
  if $x^* = \emptyset$ **then**
    3.a return BLP (4.1)-(4.3) is infeasible.
  else
    3.b return (x^*, y^*, ζ^*) is an optimal solution to the BLP (4.1)-(4.3).
  end if
else
  3.c select a node k from $\{1, ..., N\}$, set $\hat{u} = U^k$, $\hat{v} = V^k$, $\hat{x} = X^k$, $\hat{y} = Y^k$, $\hat{\lambda} = \Lambda^k$, and $\hat{z} = Z^k$. Discard node k, reorder the remaining nodes from 1 to $N - 1$, reduce N by 1, and go to Step 4.
end if

Step 4:

if $(\hat{x}, \hat{y}) = \emptyset$ **then**
  Solve $\mathcal{P}(\hat{u})$.
  if $\mathcal{P}(\hat{u})$ is infeasible **then**
    4.a go to Step 3.
  else
    Let $(x^{\mathrm{R}}, y^{\mathrm{R}})$ denote an optimal solution and ζ^{R} the optimal objective value of $\mathcal{P}(\hat{u})$. Set $\lambda^{\mathrm{R}} = \hat{\lambda}$.
  end if
else if $\hat{\lambda} = \emptyset$ **then**
  Solve $\mathcal{D}(\hat{v})$.
  if $\mathcal{D}(\hat{v})$ is infeasible **then**
    4.a go to Step 3.
  else
    Let λ^{R} denote an optimal solution of $\mathcal{D}(\hat{v})$. Set $x^{\mathrm{R}} = \hat{x}$, $y^{\mathrm{R}} = \hat{y}$, and $\zeta^{\mathrm{R}} = \hat{z}$.
  end if
end if
if $\zeta^{\mathrm{R}} \leq \zeta^*$ **then**
  4.b go to Step 3.
else if $(b_2 - A_2 x^{\mathrm{R}} - B_2 y^{\mathrm{R}}) \perp \lambda^{\mathrm{R}}$ **then**
  4.c update $x^* = x^{\mathrm{R}}$, $y^* = y^{\mathrm{R}}$, and $\zeta^* = \zeta^{\mathrm{R}}$ and go to Step 3.
else
  4.d go to Step 2.
end if
end function

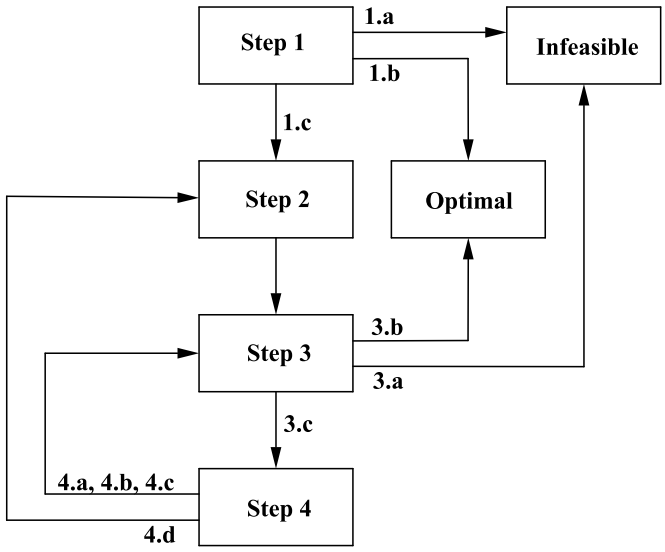


Figure 4.1: Diagram of the new branch and bound algorithm

At Step 1, the relaxation problem $\mathcal{R}(\infty, \infty)$ is solved to check if the BLP is infeasible (1.a), or has an optimal solution. In the case that the relaxation problem has an optimal solution, a check on complementarity constraints made to find a potentially bilevel optimal solution. Confirmation indicates that the optimal solution of the BLP problem is found (1.b). Alternatively, if the complementarity constraints are not satisfied (1.c), two new nodes are created at Step 2. At this step, the term with the largest product of $(b_2 - A_2x^{\mathrm{R}} - B_2y^{\mathrm{R}})_i$ and λ_i^{R} is used to define the branching variable. At Step 3, the backtracking is performed. If there is no live node, two cases can happen: if no optimal solution has been found yet (3.a), then the BLP is infeasible; otherwise, the incumbent solution is reported as the optimal solution of BLP problem (3.b). Alternatively, if there are live nodes (3.c), a node k is selected to be solved at Step 4.

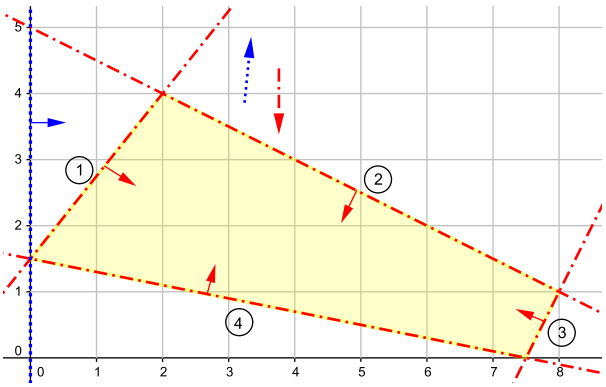
Step 4 is designed to find a potentially bilevel feasible solution. If the node makes $(b_2 - A_2x - B_2y)_i = 0$, then the problem $\mathcal{P}(u)$ is solved to find a solution (x, y), and the solution of the parent node is used for λ. Whereas, if the node makes $\lambda_i = 0$, then the problem $\mathcal{D}(v)$ is solved to find λ, and (x, y) are taken from the solution of the parent node. If $\mathcal{P}(u)$ or $\mathcal{D}(v)$ is infeasible (4.a), or the objective value of the relaxation problem is worse than the incumbent solution (4.b), the backtracking is accomplished at Step 3. If the solution satisfies

the complementarity constraints (4.c), the incumbent solution is updated, the algorithm goes to Step 3, and backtracks. When the solution of the relaxation problem is not fathomed either due to infeasibility or being worse than the incumbent solution or satisfying complementary slackness conditions, two new nodes is created at Step 2 (4.d). The algorithm terminates either at Step 1 when the root node is infeasible (1.a) or has a feasible bilevel solution (1.b), or at Step 3, when there is no live node (3.a and 3.b).

Since this algorithm is equivalent to the original branch and bound algorithm, it guarantees to terminate with global optimal solution. This design is valid and correct because of the branch and bound structure and branching method on complementarity constraints.

To illustrate how the algorithm works, a simple example from Moore and Bard (1990) is solved. Figure 4.2 shows the branch and Bound tree of this example and Table 4.2 summarizes the solution at each iteration.

$$
\begin{aligned}
\max_{x,y} \quad & x + 10y \\
\text{s. t.} \quad & -x \leq 0 \\
& \max_{y} \quad -y \\
& \text{s. t.} \quad -5x + 4y \leq 6 \\
& \qquad x + 2y \leq 10 \\
& \qquad 2x - y \leq 15 \\
& \qquad -2x - 10y \leq -15
\end{aligned}
$$



For node 0, we solve the problem $\mathcal{R}(\infty, \infty)$; the solution would be $(x^0, y^0) = (2, 4)$ and $\lambda^0 = (0, 0, 1, 0)$ with $\zeta^0 = 42$. Since the solution does not satisfy complementarity constraints, we branch on the third constraint of the lower-level problem. Therefore, we impose $u_3 = 0$ in node 1 and $v_3 = 0$ in node 2. We need to solve the problem $\mathcal{P}(\hat{u})$ for node 1 to obtain $(x^1, y^1) = (8, 1)$ with $\zeta^1 = 18$. In addition, λ^1 is taken from node 0 directly, thus, $\lambda^1 = (0, 0, 1, 0)$. The solution (x^1, y^1, λ^1) satisfies complementarity constraints. Therefore, the incumbent solution is updated, and node 1 is pruned. For node 2, we solve the problem $\mathcal{D}(\hat{v})$ to obtain $\lambda^2 = (0, 0, 0, 0.1)$.

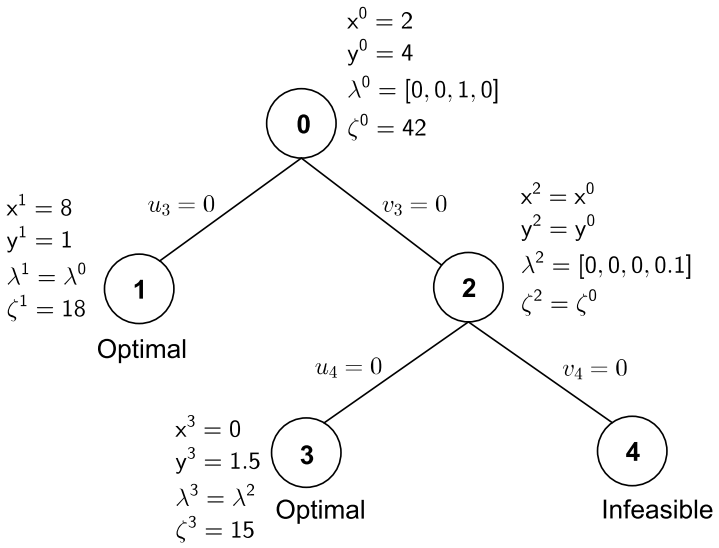


Figure 4.2: The Branch and Bound tree of the example

Variables (x^2, y^2) are taken from node 0 directly, hence $(x^2, y^2) = (2, 4)$ with $\zeta^2 = 42$. Since complementarity constraints of node 2 are not satisfied, we need to branch on the fourth constraint of the lower-level problem in node 2 to make two new nodes 3 and 4. We will have $u_4 = 0$ in node 3 and $v_4 = 0$ in node 4. Then again, we solve the problem $\mathcal{P}(\hat{u})$ for node 3 to obtain $(x^3, y^3) = (0, 1.5)$ with $\zeta^3 = 15$ and take variable $\lambda^3 = (0, 0, 0, 0.1)$ from node 2 directly. The solution of node 3 is worse than the incumbent solution, which is 18, thus it is pruned. For node 4, we solve the problem $\mathcal{D}(\hat{v})$. Since it is infeasible, and there is no more node to be solved, the incumbent solution is reported as the optimal solution. Therefore, the optimal solution is $(x^*, y^*) = (8, 1)$ with $\zeta^* = 18$. Table 4.2 summarizes the solution of nodes for the provided example.

4.3.1 An alternative objective function for the problem $\mathcal{D}(v)$

In this section, we suggest an alternative formulation of the problem $\mathcal{D}(v)$, which may help reduce the number of iterations to improve the efficiency of the algorithm. If the branch makes $v_i = 0$, we take the solution (x, y) from the parent node, thus, we know which constraints of $A_2x + B_2y \leq b_2$ are not binding before solving the problem $\mathcal{D}(v)$. Therefore, the corresponding variables λ to non-binding constraints need to be zero to satisfy complementarity constraints.

Table 4.2: The solution of nodes by implementing the new algorithm

Node number	Parent node	$\hat{u}$	$\hat{v}$	Problem	Solution	Information taken from parent node
0	–	$\begin{bmatrix}\infty\\ \infty\\ \infty\\ \infty\end{bmatrix}$	$\begin{bmatrix}\infty\\ \infty\\ \infty\\ \infty\end{bmatrix}$	$\mathcal{R}(\infty,\infty)$	$(x^0,y^0)=(2,4)$ $\lambda^0=[0,0,1,0]$ $\zeta^0=42$	–
1	0	$\begin{bmatrix}\infty\\ \infty\\ 0\\ \infty\end{bmatrix}$	$\begin{bmatrix}\infty\\ \infty\\ \infty\\ \infty\end{bmatrix}$	$\mathcal{P}(\hat{u})$	$(x^1,y^1)=(8,1)$ $\zeta^1=18$	$\lambda^1=[0,0,1,0]$
2	1	$\begin{bmatrix}\infty\\ \infty\\ \infty\\ \infty\end{bmatrix}$	$\begin{bmatrix}\infty\\ \infty\\ 0\\ \infty\end{bmatrix}$	$\mathcal{D}(\hat{v})$	$\lambda^2=[0,0,0,0.1]$	$(x^2,y^2)=(2,4)$ $\zeta^2=42$
3	2	$\begin{bmatrix}\infty\\ \infty\\ \infty\\ 0\end{bmatrix}$	$\begin{bmatrix}\infty\\ \infty\\ 0\\ \infty\end{bmatrix}$	$\mathcal{P}(\hat{u})$	$(x^3,y^3)=(0,1.5)$ $\zeta^3=15$	$\lambda^3=[0,0,0,0.1]$
4	2	$\begin{bmatrix}\infty\\ \infty\\ \infty\\ \infty\end{bmatrix}$	$\begin{bmatrix}\infty\\ \infty\\ 0\\ 0\end{bmatrix}$	$\mathcal{D}(\hat{v})$	Infeasible	–

Accordingly, we can decrease the number of iterations by changing the objective function of problem $\mathcal{D}(v)$ to minimizing the summation of corresponding λ variables. For example, if the indices of non-binding constraints are $[2,4,5]$, then the objective function of problem $\mathcal{D}(v)$ will be: $\min \lambda_2+\lambda_4+\lambda_5$. An alternative problem $\mathcal{D}(v)$ is as follows:

$$\min_{\lambda} \quad \sum_{i\in I}\lambda_i \tag{4.52}$$

$$\text{s.t.} \quad 0\le\lambda\le v \tag{4.53}$$

$$B_2^{\top}\lambda=d_2 \tag{4.54}$$

where I is the set of non-binding constraints of the lower level problem in the parent node.

4.4 Numerical results

To investigate the computational performance of the new branch and bound algorithm we use 100 randomly generated bilevel linear programming problems with different dimensions. There are ten groups of instances of different size, each one consists of ten instances with the same dimensions. All matrices and vectors are integers randomly generated from uniform distribution within a range. Constraints coefficients are within $[0, 10]$, objective functions coefficients are within $[-50, 50]$, the right-hand side of the upper-level problem is within $[30, 130]$, and the right-hand-side of the lower-level problem is within $[10, 110]$.

We compare the results of the original type of new algorithm denoted by Alg^{I}, the new algorithm with an alternative objective function denoted by Alg^{II}, and the Branch and Bound algorithm represented by BB. Computational experiments were executed on a desktop computer with a 3.6 GHz CPU and a 16GB ram. All algorithms were implemented in MATLAB and CPLEX solver. Table 4.3 reports the computation times which are three average values over ten instances in each group from top to bottom: times needed to find the first bilevel feasible solution, to obtain the optimal solution, and to confirm its global optimality. In addition, all three algorithms were run by using different searching strategies: depth first, breadth first, and largest-z first searching strategies. The two last rows in the table report the average times over all 100 instances and total times of solving all instances.

We make the following observations from the results summarized in Table 4.3. First, Alg^{II} is on average 43%, 53%, and 47% faster than branch and bound by applying depth first, breadth first, and largest-z first strategies, respectively. The total time reported in the last row of Table 4.3 is the summation of times to find the optimal solution of 100 instances. For example, if the largest-z strategy is used, the total time of solving 100 instances for BB is around two hours, for Alg^{I} is about one hour, and Alg^{II} can solve them in around 40 minutes. Second, the largest-z strategy works better than other two strategies in terms of the total time and the number of iterations to find the optimal solution. However, The depth first strategy can find the first bilevel feasible solution faster than other two searching strategies.

Table 4.3: Average computation times (in seconds) for ten groups of BLP instances.

Instance groups	Depth first			Breadth first			Largest-z first		
	BB	Alg^{I}	Alg^{II}	BB	Alg^{I}	Alg^{II}	BB	Alg^{I}	Alg^{II}
$n_1 = 10, m_1 = 24$	0	0	0	0	0	0	0	0	0
$n_2 = 10, m_2 = 14$	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0
$n_1 = 60, m_1 = 14$	0	0	0	0	0	0	0	0	0
$n_2 = 604, m_2 = 84$	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0
$n_1 = 110, m_1 = 264$	0	0	0	1	1	1	1	1	0
$n_2 = 110, m_2 = 154$	4	2	2	2	1	1	1	1	1
	4	3	2	2	1	1	1	1	1
$n_1 = 160, m_1 = 384$	1	0	1	3	2	1	4	2	2
$n_2 = 160, m_2 = 224$	25	13	6	14	8	4	4	2	2
	25	13	6	14	8	4	4	2	2
$n_1 = 210, m_1 = 504$	3	2	2	10	5	4	9	5	3
$n_2 = 210, m_2 = 294$	59	29	17	22	13	8	9	5	4
	62	31	18	24	13	9	9	5	4
$n_1 = 260, m_1 = 624$	5	3	10	36	20	23	20	11	8
$n_2 = 260, m_2 = 364$	247	122	63	111	59	35	21	11	8
	253	125	66	111	59	36	21	11	9
$n_1 = 310, m_1 = 744$	15	9	14	65	35	25	45	25	15
$n_2 = 310, m_2 = 434$	408	206	138	199	103	58	45	25	15
	430	218	140	203	105	58	52	28	15
$n_1 = 360, m_1 = 864$	29	15	29	183	97	50	142	78	31
$n_2 = 360, m_2 = 504$	1,232	560	338	541	301	156	142	78	31
	1,272	581	349	556	305	162	146	81	32
$n_1 = 410, m_1 = 984$	29	21	52	263	143	73	150	83	56
$n_2 = 410, m_2 = 574$	2,169	1,236	1,074	820	428	213	151	84	56
	2,246	1,284	1,090	850	444	218	155	87	57
$n_1 = 460, m_1 = 1104$	47	34	89	467	241	237	301	164	127
$n_2 = 460, m_2 = 644$	4,071	2,079	1,657	1,877	953	693	304	165	130
	4,179	2,136	1,683	1,901	961	695	305	165	131
Average time (s)	13	8	20	103	54	41	67	37	24
	822	425	329	359	186	117	68	37	25
	847	439	335	366	190	118	69	38	25
Total time (hh:mm)	23:32	12:12	09:19	10:10	05:16	03:17	01:56	01:04	00:42

Table 4.4 represents the average number of iterations to find the optimal solution over ten instances in each group for three algorithms and three types of searching strategies. The last row of the table summarizes the average number of iterations to obtain the optimal solution over all 100 instances. As can be seen from the table, the average number of iterations for Alg^{I} is almost the same as BB. However, Table 4.3 indicates that the Alg^{I} is faster than BB because a smaller problem is solved in each iteration. The number of iterations for Alg^{II} to find the optimal solution is on average significantly lower than branch and bound algorithm. As a result, the optimal solution is obtained in a shorter time.

Table 4.4: Average number of iterations for ten groups of BLP instances.

Instance groups	Depth first			Breadth first			Largest-z first		
	BB	Alg^{I}	Alg^{II}	BB	Alg^{I}	Alg^{II}	BB	Alg^{I}	Alg^{II}
1	18	19	16	16	17	14	11	11	11
2	100	96	52	86	86	42	48	48	36
3	386	379	289	186	186	154	74	75	67
4	988	959	445	543	532	265	161	156	121
5	1,162	1,108	614	448	455	292	191	196	125
6	2,701	2,624	1,366	1,199	1,162	703	245	235	178
7	2,790	2,743	1,742	1,231	1,191	647	334	335	179
8	4,643	4,542	2,702	2,079	2,154	1,141	593	601	240
9	5,837	5,785	4,908	2,173	2,125	1,026	455	459	296
10	7,463	7,344	5,617	2,960	3,039	2,183	579	571	447
Average	2,609	2,560	1,775	1,092	1,095	647	269	269	170

4.5 Conclusions

This paper presents a new branch and bound algorithm for solving bilevel linear programming problems, which is faster than current algorithms in the literature. The relaxation problem is subdivided into two smaller problems: $\mathcal{P}(u)$ and $\mathcal{D}(v)$. After branching a node into two new nodes, one of them solves the problem $\mathcal{P}(u)$ and another one solves the problem $\mathcal{D}(v)$. Therefore, a part of the relaxation problem is solved at each node. In addition, an alternative objective function for problem $\mathcal{D}(v)$ is introduced to also reduce the number of iterations. Therefore, it shortens the solution time even more. The results of solving the 100 randomly

generated instances with different sizes demonstrate that the proposed algorithm is faster than branch and bound algorithm in all three searching strategies.

Our future work is applying more heuristic methods to use the information of solved nodes and consequently, find the optimal solution faster. For one example, if we solve a node and it is infeasible, we may want to prune all other nodes with similar binding constraints or more. This can reduce the number of live nodes and accordingly number of iterations. In addition, most existing algorithms assume that the constraint region is compact and cannot solve the bilevel problem when the relaxation problem is unbounded. If the relaxation problem is unbounded, the bilevel solution can be either unbounded, infeasible or optimal. Therefore, one more future research direction is addressing the unbounded case in the algorithm.

BIBLIOGRAPHY

Aiyoshi, E. and Shimizu, K. (1984). A solution method for the static constrained stackelberg problem via penalty method. *IEEE Transactions on Automatic Control*, 29(12):1111–1114.

Bard, J. F. (1983). An efficient point algorithm for a linear two-stage optimization problem. *Operations Research*, 31(4):670–684.

Bard, J. F. (1991). Some properties of the bilevel programming problem. *Journal of Optimization Theory and Applications*, 68(2):371–378.

Bard, J. F. (2013). *Practical bilevel optimization: algorithms and applications*, volume 30. Springer Science & Business Media.

Bard, J. F. and Moore, J. T. (1990). A branch and bound algorithm for the bilevel programming problem. *SIAM Journal on Scientific and Statistical Computing*, 11(2):281–292.

Ben-Ayed, O. and Blair, C. E. (1990). Computational difficulties of bilevel linear programming. *Operations Research*, 38(3):556–560.

Bialas, W. and Karwan, M. (1982). On two-level optimization. *IEEE Transactions on Automatic Control*, 27(1):211–214.

Calvete, H. I., Gale, C., and Mateo, P. M. (2008). A new approach for solving linear bilevel problems using genetic algorithms. *European Journal of Operational Research*, 188(1):14–28.

Fortuny-Amat, J. and McCarl, B. (1981). A representation and economic interpretation of a two-level programming problem. *The Journal of the Operational Research Society*, pages 783–792.

Gendreau, M., Marcotte, P., and Savard, G. (1996). A hybrid tabu-ascent algorithm for the linear bilevel programming problem. *Journal of Global Optimization*, 8(3):217–233.

Hansen, P., Jaumard, B., and Savard, G. (1992). New branch-and-bound rules for linear bilevel programming. *SIAM Journal on Scientific and Statistical Computing*, 13(5):1194–1217.

Hejazi, S. R., Memariani, A., Jahanshahloo, G., and Sepehri, M. M. (2002). Linear bilevel programming solution by genetic algorithm. *Computers & Operations Research*, 29(13):1913–1925.

Hu, J., Mitchell, J. E., Pang, J.-S., Bennett, K. P., and Kunapuli, G. (2008). On the global solution of linear programs with linear complementarity constraints. *SIAM Journal on Optimization*, 19(1):445–471.

Kuo, R. J. and Huang, C. C. (2009). Application of particle swarm optimization algorithm for solving bi-level linear programming problem. *Computers & Mathematics with Applications*, 58(4):678–685.

Mathieu, R., Pittard, L., and Anandalingam, G. (1994). Genetic algorithm based approach to bi-level linear programming. *Revue française d'automatique, d'informatique et de recherche opérationnelle. Recherche opérationnelle*, 28(1):1–21.

Moore, J. T. and Bard, J. F. (1990). The mixed integer linear bilevel programming problem. *Operations research*, 38(5):911–921.

Oduguwa, V. and Roy, R. (2002). Bi-level optimisation using genetic algorithm. In *2002 IEEE International Conference on Artificial Intelligence Systems (ICAIS02)*, pages 322–327. IEEE.

Osman, M., El-Wahed, W. A., Shafei, M. M., and Wahab, H. B. (2009). A solution methodology of bi-level linear programming based on genetic algorithm. *Journal of Mathematics and Statistics*, 5(4):352.

Rajesh, J., Gupta, K., Kusumakar, H. S., Jayaraman, V. K., and Kulkarni, B. D. (2003). A tabu search based approach for solving a class of bilevel programming problems in chemical engineering. *Journal of Heuristics*, 9(4):307–319.

Sahin, K. H. and Ciric, A. R. (1998). A dual temperature simulated annealing approach for solving bilevel programming problems. *Computers & Chemical Engineering*, 23(1):11–25.

Vicente, L., Savard, G., and Júdice, J. (1994). Descent approaches for quadratic bilevel programming. *Journal of Optimization Theory and Applications*, 81(2):379–399.

Wang, G.-M., Wan, Z.-P., Wang, X.-J., and Chen, Y.-L. (2005). Genetic algorithms for solving linear bilevel programming. In *Sixth International Conference on Parallel and Distributed Computing Applications and Technologies (PDCAT'05)*, pages 920–924. IEEE.

White, D. J. and Anandalingam, G. (1993). A penalty function approach for solving bi-level linear programs. *Journal of Global Optimization*, 3(4):397–419.

Yin, Y. (2000). Genetic-algorithms-based approach for bilevel programming models. *Journal of Transportation Engineering*, 126(2):115–120.

CHAPTER 5. SUMMARY AND DISCUSSION

The presented research includes three papers, which make significant contributions to the area of multi-level programming problems. These contributions are discussed in this chapter.

In Chapter 2, we introduced a model to investigate the potential competition for biomass between two major renewable energy policies, RPS and RFS2, as well as other interactions between them. Unlike previous research that studied these two policies separately and considered a subset of the resource and geographical dimensions, we analyzed the interactions between RPS and RFS2 on the comprehensive renewable energy portfolio. The problem is a multi-level decision-making model with the government, policy makers, biofuel producers, electricity generators, and farmers as decision makers. However, there are mainly two levels of decision makers, including policy makers and producers/generators. We assumed that renewable energy requirements, energy prices, and tax credits, which are the variables of the leader, were defined and we focused on the lower level problem to predict the renewable energy generation under interactions of RPS and RFS2 policies. Furthermore, to investigate the possible changes in the prediction, we performed a sensitivity analysis by changing the model parameters. However, a multi-level decision-making model can be developed to determine the renewable energy policy requirements, energy prices, and tax credits for future work.

In Chapter 3, we developed a model to determine the order policy of a single item in an incapacitated warehouse, where the demand and lead time were both uncertain. The warehouse orders new items to satisfy the demand over an infinite discrete time period to minimize the total cost. Since uncertain demand is observed in each period and the lead time is realized when an order arrives, it is a multi-stage decision-making problem. We reformulated the multi-stage decision-making model as a two-stage tri-level optimization model. The first stage refers to the first period of the planning horizon, and all remaining periods are considered as the second stage. Therefore, it is assumed that after making the first stage decisions, all uncertain parameters of

the second stage will be observable, thus, the second stage becomes a deterministic problem. In the proposed tri-level model, the upper-level makes the first stage decisions, the middle-level realizes the uncertain parameters such that the total cost is maximized to identify the worst case scenario, and the lower-level makes the second stage decisions. We ran the model in a rolling planning horizon, in which only the first stage decisions are actually implemented. The results of the model were compared to the deterministic model with optimistic, moderate, and pessimistic approaches and a sensitivity analysis on cost parameters were done to examine the robustness of the order policy. The results show that the tri-level model determines the worst-case scenario effectively depending on different cost parameters.

In Chapter 4, we proposed an exact algorithm for solving the bilevel linear programming problem. This type of problem has many important application areas, such as economics, transportation, and business. However, many researchers have discussed the difficulties of developing efficient algorithms to solve it and proved that it is NP-hard. We developed an algorithm to solve BLP problems faster than current algorithms. We reformulated the two-level model as a single-level model with linear complementarity constraints by applying "Karush-Kuhn-Tucker" conditions. The reformulated problem can be solved by different algorithms, such as branch and bound and Benders decomposition. We used a branch and bound scheme to meet the complementarity slackness conditions. The relaxation problem, which is solved in each node, is obtained by relaxing the complementarity conditions. At each iteration, a check on complementarity slackness is made to see if it is satisfied. If so, the solution is a bilevel feasible solution; if not, branching is performed to examine all combinations of complementarity conditions. In our new algorithm, we subdivided the relaxation problem into two smaller problems to save time in each node. The integration of these two problems is equivalent to the relaxation problem. In addition, we introduced an alternative objective function for the relaxation problem to decrease the number of iteration. The results of solving 100 examples indicate that the new algorithm is faster than the original branch and bound method.

Future work toward the first paper can be extending the model to a multi-level programming problem, which includes the strategic behavior of policy makers and investors in the renewable energy markets. This model would need a entirely different modeling approach with

limitations of its own. Furthermore, considering hydropower in different RPS legislations is another extension to the model. As for the second paper, we developed the inventory control model for a single item. An extension can consider multiple items and transportation cost in the model because some items can be shipped together as a batch. Therefore, we need a model to decide what items can be grouped regarding their order cycles. Our future work for the third paper is applying heuristic methods to reduce the number of iterations by using the solution of solved nodes. For example, if a node is infeasible, we can prune all nodes with similar binding constraints or more. For another example, if the problem $\mathcal{D}(v)$ is feasible for a node, its solution is also feasible for other nodes with similar and less zero entities in vector v; thus, we can consider this solution for those nodes without solving them. In addition, a few existing algorithms can solve the bilevel problem without assuming that the constraint region is compact. One more future research direction is addressing the unbounded case in the algorithm.