OPTIMAL DESIGN OF THE ANNUAL INFLUENZA VACCINE

by

Osman Y. Özaltın

B.S., Boğaziçi University, 2005

M.S. in Industrial Engineering, University of Pittsburgh, 2007

Submitted to the Graduate Faculty of

the Swanson School of Engineering in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2011

UNIVERSITY OF PITTSBURGH

SWANSON SCHOOL OF ENGINEERING

This dissertation was presented

by

Osman Y. Özaltın

It was defended on

July 29, 2011

and approved by

Andrew J. Schaefer, Professor, Department of Industrial Engineering

Oleg A. Prokopyev, Assistant Professor, Department of Industrial Engineering

Michael Trick, Professor, Tepper School of Business, Carnegie-Mellon University

Juan P. Vielma, Assistant Professor, Assistant Professor, Department of Industrial Engineering

Dissertation Directors: Andrew J. Schaefer, Professor, Department of Industrial Engineering,

Oleg A. Prokopyev, Assistant Professor, Department of Industrial Engineering

OPTIMAL DESIGN OF THE ANNUAL INFLUENZA VACCINE

Osman Y. Özaltın, PhD

University of Pittsburgh, 2011

Seasonal influenza is a major public health concern, and the first line of defense is the flu shot. Antigenic drifts and the high rate of influenza transmission require annual updates to the flu shot composition. The World Health Organization recommends which flu strains to include in the annual vaccine based on surveillance and epidemiological analysis. There are two critical decisions regarding the flu shot design. One is its composition; currently, three strains constitute the flu shot, and they influence vaccine effectiveness. Another critical decision is the timing of the composition decisions, which affects the flu shot production. Both of these decisions have to be made under uncertainty many months before the flu season starts.

We quantify the trade offs involved through multi-stage stochastic mixed-integer programs that determine the optimal flu shot composition and its timing in a stochastic and dynamic environment. Our first model takes the view of a social planner, and optimizes strain selections based on a production plan that is provided by the flu shot manufacturers. It also incorporates risk-sensitivity through mean-risk models. Our second model relaxes the exogenous production planning assumption and, hence, provides a more accurate representation of the hierarchical decision mechanism between a social planner, who selects the flu shot strains, and the manufacturers, who make the flu shot available. We derive structural properties of both models, and calibrate them using publicly available data.

The flu shot strains are updated based on clinical, virological and immunological surveillance. In the virological surveillance, hemagglutinin inhibition assays are used to identify antigenic properties of the influenza viruses. However, this serology assay is labor-intensive

and time-consuming. As an alternative, pairwise amino acid sequence comparison of influenza strains is used in statistical learning models to identify positions that cause antigenic variety. The performance of these models is evaluated by cross validation. In Chapter 5, we formulate cross validation as a bilevel program where an upper-level program chooses the model variables to minimize the out-of-sample error, and lower-level problems in each fold optimize in-sample errors according to their training data set by selecting the regression coefficients of the chosen model variables. We provide an extensive computational study using clinical data, and identify amino acid positions that significantly contribute to antigenic variety of influenza strains.

Keywords: Operations research, mixed-integer programming, multistage stochastic programming, bilevel programming, column generation, branch-and-price, oligopolistic market, influenza vaccine design.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENTS

to my family

This work would never had been possible without my wonderful family: my parents Sare and Orhan, my brother Ferhat, and his wife Tuğba. They have been always there for me when I needed. Their endless love, encouragement, and unconditional support have always been an inspiration and strength to me. I have added another wonderful family to my life by marrying my wife: my parents-in-law Müjgan and Kazım, and my brothers-in-law Kutalmış Emre and Ahmet Cevdet, to whom I owe a lot. Your prayers are all accepted. Finally, the person I owe more thanks to than all of the previous people combined: my enduring, caring, encouraging, and loving wife Nur Özge, my invaluable gift from Allah.

1.0 INTRODUCTION

Health care is the largest industry in the United States and continues to grow. National health care expenditures grew 4.0% to $2.5 trillion in 2009, or $8,086 per person, and accounted for 17.6% of the Gross Domestic Product (GDP). Furthermore, growth in spending is projected to average 6.1% annually over the period 2009 through 2019, reaching 19.3% of GDP by 2019 [41]. Currently, among 193 member countries of the World Health Organization (WHO), total health expenditure of the U.S. is the highest both on per-capita basis and as percent of the GDP [190]. However, the WHO ranked the U.S. health system as 37^{th} in overall performance [189]. This inefficiency of the U.S. health care system have resulted in a growing research interest in medical decision making over the last few decades.

Policy level planning and control decisions for health care services are made in the design phase, whereas operational decisions are confronted daily and often in high volumes. Operations Research (OR) techniques have been applied to a wide range of decision problems in both levels, which have significant effects on cost, quality and patient satisfaction. Pierskalla and Brailer [151] have classified those problems into three main groups. In the first group, examples of system design and planning applications include treatment resource planning [158], bed capacity planning [82], regional design for organ procurement equity and efficiency [108, 172], developing pediatric immunization schedules [167], strain selection for the annual influenza vaccine [45, 109, 191], and control of infectious diseases [16, 18, 125, 157]. In the second group, examples of health care operations management are patient scheduling [149], workforce planning and scheduling [99, 139], and emergency room or operating room scheduling [55]. In the last group, examples of medical decision making comprise screening for disease [43, 131], cancer treatment planning [60, 120, 121, 136, 153, 159, 168], and the optimal timing of organ transplants [4, 5, 6, 91, 94, 163].

This dissertation focuses on the annual influenza vaccine, hereafter the *flu shot*. Influenza is a highly contagious, acute, respiratory viral disease. Seasonal flu epidemics impact 5–15% of the world's population, resulting in 3–5 million cases of severe illnesses and up to 500,000 deaths annually [185]. Pneumonia, a common consequence of influenza, is the 8$^{\text{th}}$-leading cause of death in the United States [115]. In a typical year, 5–20% of the United States' population gets the flu, of whom 200,000 are hospitalized and 36,000 die [39]. Higher risk population groups include older adults, those with chronic medical conditions, pregnant women and young children [35]. Seasonal flu has significant economic impacts such as direct treatment costs, work absenteeism and reduced work productivity. These costs are estimated to be \$12–14 billion annually in the United States [27].

The flu shot is the first line of defense against seasonal epidemics [39]. However, circulating influenza viruses frequently mutate in response to antibody pressure, and the flu shot provides protection only if its composition is antigenically similar to those strains. Therefore, the vaccine strains have to be updated frequently. There are two critical decisions regarding the flu shot design. One is its *composition*, which determines the vaccine effectiveness. As recently as the 2007–08 flu season, a poor match between the selected strains and the ones that actually emerged in the season reduced the vaccine effectiveness by half. Figure 1.1 depicts the prevalence of the flu shot strains as well as the other strains emerged but were not included in the vaccine in the 2007–08 season.

The other critical decision is the *timing* of the composition decisions, which affects the flu shot availability. The following quote is from a flu shot manufacturer [68]:
"The timing of the strain selection is also important. We have a limited production time due to the necessity of distributing and administering vaccine prior to the influenza season."

The flu shot supply and its distribution cannot be guaranteed in any year, as the composition and timing decisions have to be made under uncertainty at least six months before the epidemic [74]. If the strains are selected too early with insufficient surveillance data, there is a greater likelihood that there will be a mismatch. However, delaying the decisions may adversely impact production lead times, resulting in delays or shortages in the flu shot supply. As can be seen in Figure 1.2, repeated disruptions in the manufacture and distribution of the flu shot have occurred in the last decade [28, 29].

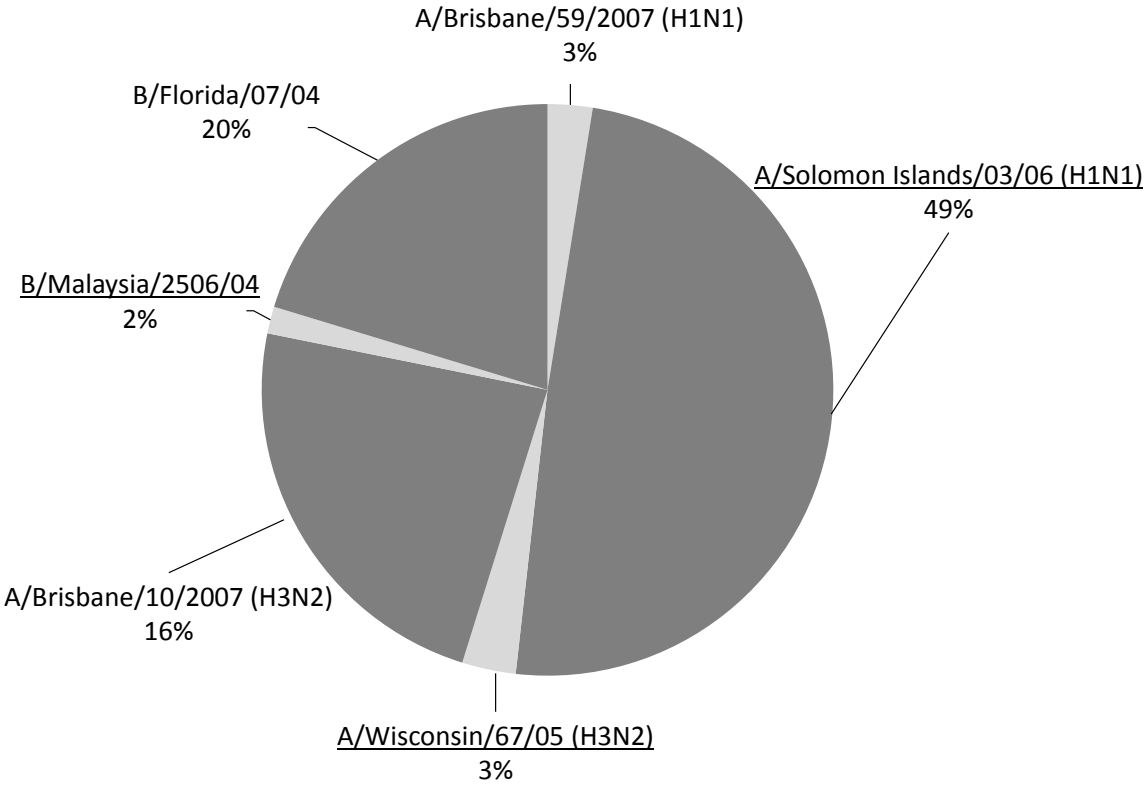


Figure 1.1: Prevalence of the circulating influenza strains in the 2007-08 season. Two of the three vaccine strains failed to match the flu strains that actually emerged in the season [33]. Strains whose names are underlined are chosen for the flu shot.

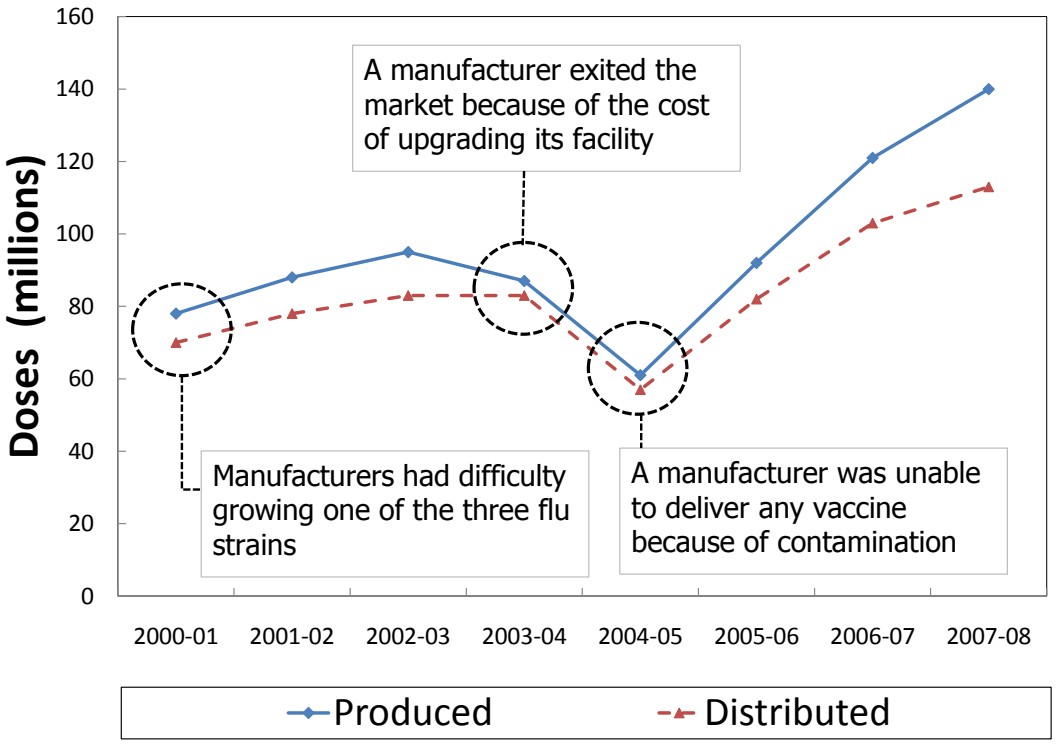


Figure 1.2: Disruptions in the manufacture and distribution of the flu shot in the last decade [90].

1.1 THE FLU VIRUS

There are three serotypes of the flu virus: influenza A, B and C. Influenza A and B viruses are mainly responsible for seasonal flu epidemics, whereas influenza C viruses are less common and usually cause mild upper respiratory illnesses [185].

Influenza A viruses are subtyped on the basis of their two surface proteins: hemagglutinin and neuraminidase. Influenza A subtypes and influenza B viruses are further classified into *strains* based on their antigenic properties. Antigenic variations of the strains emerge according to two mechanisms: *antigenic drifts*, which apply to both A and B viruses and occur as a result of point mutations in the surface proteins; and *antigenic shifts*, which apply to only A viruses and occur as a result of replacement of surface proteins with novel ones [37].

Antigenic shifts are particularly dangerous, as they can lead to a flu pandemic, e.g. the Asian Flu Pandemic of 1957 that killed two million people worldwide [47]. Seasonal flu epidemics are caused by antigenic drifts. Although a model similar to the one described herein may apply to pandemic flu, current policy is to design pandemic flu vaccine only after the pandemic has emerged [88]; we therefore restrict our consideration to seasonal flu epidemics.

1.2 THE FLU SHOT DESIGN

Currently, the flu shot contains inactivated strains of two influenza A subtypes (H3N2 and H1N1) and one influenza B virus. However, antigenic drifts and the high rate of influenza transmission require frequent changes in the flu shot composition. Between 1980 and 2010, 16 changes were made for the A/H3N2 strain, 13 for the A/H1N1 strain, and 15 for the B strain [184].

The World Health Organization (WHO) coordinates an international influenza surveillance network, which is responsible for making detailed analysis of circulating strains. Based on surveillance data, the WHO annually recommends which strains to include in the flu shot. Currently, recommendations are made in February-March for the Northern hemisphere, and

in August-September for the Southern hemisphere. Although seasonal flu is a global health concern, we focus on the United States where the Vaccines and Related Biological Products Advisory Committee of Food and Drug Administration (FDA) (hereafter "the Committee") makes the final decision about the flu shot composition soon after the WHO recommendations. In the United States, Center for Disease Control (CDC) is responsible for influenza surveillance.

Antigenically different flu strains may co–circulate in one season, which brings an additional challenge for determining the flu shot composition. The flu shot provides protection against antigenically similar strains, a phenomenon known as *cross–effective immunity* [36]. The Committee uses Hemagglutination Inhibition (HI) assays to measure cross–effectiveness among the strains. A large amount of data is available from surveillance; however, in current practice, the Committee does not appear to have analytical tools to utilize these data. This fact is verified by a quote from a Committee member [67]:
"... maybe we should select the non-dominant strain for this past year in anticipation of the fact that it may be the dominant strain next year. But you are right, it is a guessing game."

1.3 THE FLU SHOT PRODUCTION

Currently, six manufacturers provide the flu shot to the United States market [40]. Representatives of these manufacturers attend the Committee meetings to discuss the production capability of candidate strains [69]. However, once the strains are selected, manufacturers make their own production plans to maximize profits.

The flu shot production has many critical and time-sensitive steps. Strain production in fertilized chicken eggs starts right after the high-growth seed strains become ready. Harvesting, purification, testing, and packaging occur from June to November. Finally, the FDA must approve each lot prior to shipment. The amount of the flu shot produced is limited by the lowest strain production quantity, as the strains are combined together to compose the flu shot.

The production yield ratio of strains is uncertain, and difficulties growing a strain might result in reductions of the overall flu shot supply, as was the case in the 2000-01 flu season [73]. Therefore, manufacturability of a new flu shot strain has been a primary concern as stated by a flu shot manufacturer [67]:

"... we ask you here to consider not only the antigenic match, but the ability of the strains you select to enable us to produce sufficient vaccine for the marketplace. We need seed viruses, especially the high growth reassortants."

Furthermore, the number of chicken eggs used in strain production must be chosen a priori, which makes it very hard to increase the supply subsequently. Therefore, we assume that the production decisions are one-time and irreversible.

1.4 PROBLEM STATEMENT AND CONTRIBUTIONS

A large amount of data is available from clinical, virological and immunological influenza surveillance [31]; however, the Committee does not have analytical tools to exploit those data. The contribution of this dissertation is two-fold. First, we formulate mathematical programming models for the annual flu shot design problem. Then, using these models we address pressing public policy issues regarding the allocation of scarce flu prevention resources.

The main trade off in the optimization of the strain selection decisions arises between the composition of the annual flu shot and the timing of its production. On one hand, deferring the strain selection decisions increase the likelihood of delays and shortages in the flu shot supply. On the other hand, as the Committee waits, more information becomes available from the surveillance mechanism that reduces the likelihood of mismatch between the flu shot strains and those actually emerge in the season. In solving this optimization problem, we are particularly interested in the objective of maximizing the societal benefit, which is measured by the cost of infections averted by the vaccinated people minus the cost of infections incurred by unvaccinated people. We leave the consideration of other objectives for future research.

Table 1.1: Candidate flu shot strains for the 2008-09 season [67].

Type	Strain Name	Yield Ratio
A/H1N1	A/New Caledonia/20/99†	High
	A/Solomon Islands/3/06	Moderate
	A/Brisbane/59/07	Moderate/High
	A/South Dakota/6/07	Low/Moderate
A/H3N2	A/Wisconsin/67/2005†	Moderate/High
	A/Brisbane/10/2007	Moderate/High
B	B/Malaysia/2506/2004†	Moderate
	B/Florida/04/2006	Low/Moderate
	B/Brisbane/03/2007	Moderate

†Strain included in 2007-08 influenza vaccine [67].

Different aspects of the flu shot design have been studied in the literature as reviewed in Section 2.4. A common assumption is that the composition decisions are independent among different strain categories [45, 109, 191]. However, in practice, vaccine stains are grown separately and then combined together to compose the flu shot. As previously noted, the total flu shot production is limited by the "least productive strain." Therefore, unanticipated problems in growing strain from a particular category (i.e. A/H3N2, A/H1N1 or B) would result in reductions in the overall flu shot supply. Another common assumption in the literature is that for each category the most prevalent circulating strain is the only alternative for the current vaccine strain in that category [45, 109]. However, as can be seen in Table 1.1, the Committee considers multiple candidate strains per category, as a strain, which is not necessarily the most prevalent one, might be favorable based on its manufacturability and cross-effectiveness.

Our models relax both of the aforementioned assumptions in the literature by considering multiple candidate strains (not only the most prevalent ones) and all three strain categories

simultaneously. We also consider risk sensitivity using mean-risk objective functions. We address multiple policy questions after carefully calibrating our model using publicly available CDC and FDA data.

The flu shot provides immunity protection if its composition is antigenically similar to the circulating strains. In the virological surveillance, hemagglutinin inhibition (HI) assays are used to identify antigenic properties of the influenza viruses. However, this serology assay is labor-intensive and time-consuming. Alternatively, modeling the antigenic distance among influenza strains using amino acid sequence analysis can provide a rapid indication of the likelihood that the current vaccine will protect against a recently emerged strain, and also facilitates the study of the virus' evolution in response to antibody pressure [23, 171].

Statistical learning models are frequently used in sequence analysis [123, 126]. Typically, the performance of those models are evaluated by cross validation. We formulate cross validation as a bilevel program. This approach allows for optimizing the cross validation outcome and offers modeling flexibility when considering multiple statistical learning goals.

2.0 LITERATURE REVIEW

In this chapter, we review the literature related to the problems and methodologies discussed in this dissertation. Section 2.1 and Section 2.2 review multi-stage stochastic integer programming and bilevel programming, respectively. Section 2.3 summarizes the branch and price and column generation techniques. Section 2.4 presents a survey on OR studies related to the flu shot.

2.1 MULTI-STAGE STOCHASTIC INTEGER PROGRAMMING

Multi-stage stochastic programs are appropriate for modeling problems that involve finite horizon sequential decision making under uncertainty. At each time stage some portion of uncertainty is unveiled and additional information is gained about the system. The information is modeled as a discrete time stochastic process $\{\xi_t\}_{t=1}^{T}$ where T is the length of the study horizon. We assume that ξ has a finite support with discrete probability distribution $P = (p^1, \ldots, p^s)$ and the decision maker is risk–neutral. Therefore the objective function considered in this work is based on taking the expectation of outcomes over all scenarios. We incorporate risk sensitivity through mean–risk models in Section 3.2. Another assumption is that the probability distribution does not depend on optimization decisions [191]. We discuss the implications of relaxing this assumption in Section 6.2.

Decisions made at time stage t must obey the nonanticipativity property, which requires that decisions made up through time period t under those scenarios that are indistinguishable at t must be the same. This property constitutes the essential difference between the stochastic and deterministic multi–period optimization problems.

Multi–stage stochastic programs can be modeled by a scenario tree $\mathscr{T} = (\mathcal{N}, \mathcal{A})$ with T stages (see Figure 3.1). Each node m in the tree, except the root, is joined to its ancestor $a(m)$. The set of nodes at time stage t is represented by $\mathcal{N}_t$. The probability of node m is p_m where $\sum_{m \in \mathcal{N}_t} p_m = 1$ for all $t \leq T$. Moreover, $\mathscr{P}(m)$ is the unique path linking the root node to m and $\mathscr{T}(m)$ is the subtree rooted at m. Paths between the root node and the terminal nodes correspond to individual scenarios. Thus, nodes at level t represent possible scenario realizations at time stage t, and terminal nodes are also known as scenarios. Note that scenario tree notation automatically satisfies the nonanticipativity condition.

Algorithms for multi-stage stochastic mixed-integer programs do not scale well. Primal methods employ scenario tree formulation. For stochastic linear programs, the convexity of the expected recourse function motivates cutting plane algorithms [160]. Starting from the root node, primal solutions are passed down the tree where they are used to generate feasibility and optimality cuts. These cuts are passed upwards to refine the primal solutions which are again passed down the tree and so on. However, in the case of stochastic-mixed integer programs, both convexity and continuity are lost. In this case, solution techniques include enumeration [165] and branch-and-bound [3, 26] algorithms.

Dual methods employ scenario formulation where decision variables exist for every time stage in each scenario. The nonanticipativity constraints are dualized to decompose the stochastic programming problem into scenario subproblems [25, 176, 177]. Typically, there is significant duality gap, which can be addressed by progressive hedging [176, 177] or branch-and-bound techniques [8, 7, 129, 169].

Valid inequalities for mixed–integer programs are useful for tightening the formulation at the root node and speeding up the solution process at other nodes of the search tree in a branch-and-cut framework. Similarly valid inequalities for stochastic mixed–integer programs have also been recently derived [79]. We refer the reader to Guan [78] for a detailed survey of multi-stage stochastic mixed–integer programs.

2.2 BILEVEL PROGRAMMING

Bilevel programs [46, 53, 138, 147] model the hierarchical relationship between two autonomous, and possibly conflicting, decision makers: the *leader* and the *follower*. Each decision maker controls a distinct set of variables and the decisions are made sequentially according to a hierarchy: the *upper-level* decisions are made first by the leader, after which the *lower-level* decisions are made by the follower subject to the constraints which depend on the leader's decisions.

The general bilevel programming problem can be formulated as:

$$\begin{aligned} \max \quad & F(x, y), && (2.1)\\ \text{subject to} \quad & x \in \mathbb{X}, && (2.2)\\ & y \in R(x) \equiv \operatorname{argmax}\{f(x, y) \mid g(x, y) \leq 0,\ y \in \mathbb{Y}\}, && (2.3) \end{aligned}$$

where $x \in \mathbb{X} \subseteq \mathbb{R}^{n_1} \times \mathbb{Z}^{n_2}$ are the leader's (or upper-level) variables, and $y \in \mathbb{Y} \subseteq \mathbb{R}^{a_1} \times \mathbb{Z}^{a_2}$ are the follower's (or lower-level) variables. Likewise, functions $F : \mathbb{X} \times \mathbb{Y} \to \mathbb{R}^1$ and $f : \mathbb{X} \times \mathbb{Y} \to \mathbb{R}^1$ are the leader's (upper-level) and the follower's (lower-level) objective functions, respectively. The vector-valued function $g : \mathbb{X} \times \mathbb{Y} \to \mathbb{R}^{m_2}$ corresponds to the follower's (lower-level) constraints given the leader's decisions $x \in \mathbb{X} \subseteq \mathbb{R}^{n_1} \times \mathbb{Z}^{n_2}$.

There are two modeling approaches to bilevel programming [127]. In the *optimistic* case, whenever the rational reaction set is not a singleton, the follower implements a so-called *strong solution* that maximizes the objective function of the leader. That is, given the leader's decision $\bar{x} \in \mathbb{X}$, the follower's solution $y_o \in \mathbb{Y}$ satisfies $F(\bar{x}, y_o) \geq F(\bar{x}, y)\ \forall y \in R(\bar{x})$. In the *pessimistic* case, the leader assumes that whenever the follower is facing ties, he selects a so-called *weak solution* that minimizes the leader's objective function, i.e., the follower's decision $y_p \in \mathbb{Y}$ is such that $F(\bar{x}, y_p) \leq F(\bar{x}, y)\ \forall y \in R(\bar{x})$. The optimistic case might arise in a collaborative environment when both the leader and follower are allowed to cooperate, while the pessimistic case corresponds to an adversarial environment, or a risk-averse leader, who wants to minimize the worst-possible "damage" inflicted from an noncooperative follower [46].

Any mixed 0–1 programming problem can be reduced to an instance of bilevel linear programming problem [11]. Therefore, bilevel linear programming is strongly NP-hard [81]. Bard and Moore [13] and Hansen et al. [81] propose a branch and bound framework for bilevel linear programming problems that branches on the complementarity conditions of the follower's problem. An approximation approach by Júdice and Faustino [98] provides an ϵ-optimal solution. A survey by Ben-Ayed [17] provides a detailed discussion on bilevel linear programming.

Moore and Bard [140] propose a branch and bound algorithm and a set of heuristics for bilevel mixed-integer programming. Furthermore, a branch-and-cut algorithm has been recently developed by DeNegre and Ralphs [54]. Solution approaches for specific classes of bilevel mixed-integer programs where either the leader's or the follower's variables are all continuous are studied by Audet et al. [11], Dempe [52] and Wen and Yang [182]. Detailed surveys on bilevel mixed-integer programming solution techniques are provided in [12, 53, 138].

Application areas of bivel programming include revenue management [42], congestion management [22, 83], management of hazardous materials [100], management in energy sector [14], network design [132] and traffic planning [137] problems. A related broad class of problems which can be recast as a bilevel program is the *network interdiction* [85, 188], where the follower (*enemy*) attempts to maximize flow through a capacitated network, while the leader (*interdictor*) tries to minimize this maximum flow by interdicting, e.g., destroying some arcs or decreasing their capacity. Colson et al. [46] provides a comprehensive survey on bilevel programming applications.

2.3 BRANCH AND PRICE

First proposed by Ford and Fulkerson [71], column generation is used to solve linear programs that have a large number of variables. Instead of considering all columns explicitly, this technique employs a restricted master problem (RMP) that contains a subset of the columns. The algorithm follows a loop in which the RMP is solved, and the dual solution is passed

on to the pricing problem to find favorable columns. If such columns exist, a subset of them are inserted into the master problem; if not, the algorithm stops with the optimal solution. Lübbecke and Desrosiers [128] provide a detailed discussion of column generation.

Branch-and-price integrates branch-and-bound and column generation methods for solving large-scale integer programs. It is first proposed by Desrosiers et al. [58] for vehicle routing problem with time windows. At each node of the branch-and-bound tree, column generation is used to solve the LP relaxation. Branching occurs when no columns price out to enter the basis and the LP solution does not satisfy integrality conditions. Barnhart et al. [15] and Vanderbeck and Wolsey [180] present generic branch-and-price frameworks. Algorithmic efficacy is considered in Desaulniers et al. [57] and Vanderbeck [179].

2.4 STUDIES RELATED TO INFLUENZA

Deterministic epidemiological models, also known as compartmental models, are used in the study of the infectious diseases at the population scale [59, 66, 193]. These models categorize members of the population into different subgroups (compartments) according to the stages of the infectious disease, e.g. susceptible (S), infectious (I), recovered (R) [9]. Transitions among these subgroups constitute the main dynamics of the model.

Accurate representations of antibody-antigen binding are required for building realistic quantitative models of influenza spread. Perelson and Oster [150] propose a model for antibody-antigen binding. Their model represents antibodies and antigens by unique vectors in Euclidean space (the shape space). The coordinates of vectors represent generalized physio-chemical properties related to binding. The shape-space distance between an antibody and an antigen represents their affinity. Lapades and Farber [117] provide algorithms that construct explicit, quantitative coordinates for points in the shape space based on experimental data.

The flu shot protects against severe complications of the flu [92, 144]. There is a considerable research effort on the analysis of influenza vaccine cost-effectiveness for various population risk groups [65, 135, 181]. In the rest of this section, we focus on the studies about the influenza in the OR literature.

Kornish and Keeney [109] formulate the annual influenza vaccine strain selection problem as a finite-horizon optimal stopping problem. At each decision epoch, the decision maker either selects one of two candidate strains, so ending the process, or defers the selection. Analysis of this model indicates the importance of improved manufacturing techniques, as it would allow selecting the strains later when more information is available.

Deo and Corbett [56] study the impact of production yield uncertainty on the flu shot supply using a two-stage model of Cournot competition with endogenous market entry. This model extends the one presented in [109] by considering multiple manufacturers in an oligopolistic market, but does not consider the flu shot composition and its timing. A key finding of [56] is that the government should consider subsidizing research into more reliable manufacturing processes to increase the societal benefit of the flu shot.

Cho [45] builds upon both Kornish and Keeney [109] and Deo and Corbett [56] by considering the flu shot composition and production under yield uncertainty in a two-stage game model. The existence of an optimal threshold policy for when to retain the current strain, change to a new strain, or defer is shown. Moreover, policies which enhance awareness of indirect costs of infection are found to improve social welfare.

Chick et al. [44] also studies the flu shot supply chain, but focuses on contract design between a single manufacturer and a government. The manufacturer determines the production volume to maximize its profits while the government considers the costs and public health benefits of vaccination. This interaction is modeled as a sequential game. It is shown that a global social optimum can be achieved by sharing production uncertainty risks between the manufacturer and the government.

Unlike the influenza vaccine supply chain models, which make simplifying assumptions regarding the antigenic properties of the flu strains, Wu et al. [191] formulates the strain selection problem as a stochastic dynamic program using the antigenic shape–space model of Perelson and Oster [150]. In each stage, a flu epidemic that comprises only one strain is

considered, and a random walk assumption is made for the antigenic evolution. The flu shot strains are optimized for each individual based on an approximation of her antigenic history. The paper discusses the significant impediments to such personalized flu shots.

Larson [118] investigates the effectiveness of social distancing as a pandemic flu spread control method. Social contact behavior of the population is depicted by a nonhomogeneous probabilistic mixing model. A difference equation model is proposed to evaluate day-to-day evolution of the disease based on daily human contact frequencies and infection propensities of the susceptible population.

Tanner et al. [178] formulates stochastic programming models to find vaccination policies for controlling infectious disease epidemics. These models consider vaccine distribution; hence the strain selection or the characteristics of the influenza vaccine supply chain are not discussed.

3.0 OPTIMIZING THE SOCIETAL BENEFITS OF THE ANNUAL INFLUENZA VACCINE

In this chapter, we take the view of the Committee, and consider the problem of maximizing the societal benefits of the annual influenza vaccine. We formulate a multi-stage stochastic mixed-integer program that determines the optimal influenza vaccine composition and timing in a stochastic and dynamic environment. Some portion of the uncertainty around the prevalence of strains, strain production, and the severity of the season unfolds at each time stage, and for each strain category, the Committee either commits to a particular strain, or defers the decision to a later stage to wait for more surveillance data. We consider strain production levels as parameters of our model. Multiple candidate strains (not only the most prevalent ones) and all three strain categories (A/H1N1, A/H3N2, B) are included in the model simultaneously.

We assume that a production plan that is exogenously designed by the manufacturers is available. Note that a model that optimizes strain selections and production quantities simultaneously from the same perspective is not realistic, as governmental agencies, which select the strains, have little influence on the manufacturers other than approving each lot prior to distribution. This is verified by a Committee member [67]:
"I think the manufacturers know that this Committee has always given serious consideration to the availability of vaccine. And we don't make vaccine available. You [the manufacturers] make vaccine available. So that [the availability of vaccine] always enters into our decisions on what to recommend."

We exploit several structural properties, and propose a tailored branch-and-price algorithm. In addition, we incorporate risk-sensitivity through mean-risk models. We address multiple policy questions after carefully calibrating our model using publicly available CDC and FDA data. Our numerical experiments provide valuable insights for pressing policy issues. This chapter is based on Özaltın et al. [148].

3.1 MULTI-STAGE STOCHASTIC PROGRAMMING MODEL FORMULATION

Let $\mathcal{I}$ be the set of candidate strains for the annual influenza vaccine in a T-stage problem. Composition decisions are made in the first $T-1$ stages, and the uncertainties about the characteristics of the flu season are realized in the last stage. Note that the number of possible flu shot designs grows exponentially with the number of time stages. Moreover, the large number of scenarios exacerbates the situation, leading to a combinatorial explosion. Clearly, a method that requires the enumeration of all possible decisions is impractical.

We model the strain selection problem as a multi–stage stochastic mixed–integer program using a scenario tree $\mathscr{T} = (\mathcal{N}, \mathcal{A})$ as seen in Figure 3.1. We assume a finite number of scenarios whose discrete probability distribution is independent of the flu shot design [191]. An essential difference between stochastic and deterministic multi–period optimization problems is the nonanticipativity property, which requires decisions made under indistinguishable scenarios to be the same. A scenario tree formulation automatically satisfies this condition. Uncertain elements of our model include:

- e_n^i, the prevalence of strain $i \in \mathcal{I}$ in the flu season (i.e. the ratio of the number of flu infections caused by strain i to the whole population) at terminal node $n \in \mathcal{N}_T$,
- β_n^i, the production yield ratio of strain $i \in \mathcal{I}$ (i.e. the ratio of the strain production yield in doses to the number of chicken eggs used for that strain) at terminal node $n \in \mathcal{N}_T$,
- R_m^i, the production quantity of strain $i \in \mathcal{I}$ (i.e. the number of chicken eggs used for the production of strain i) when it is selected at node $m \in \mathcal{N}_{t<T}$, where $\mathcal{N}_{t<T}$ denotes $\bigcup_{t<T} \mathcal{N}_t$, and

- c_n, the average cost of a flu case; d_n, the flu vaccine demand; and q_n, the unit shortage cost per person at terminal node $n \in \mathcal{N}_T$.

We assume that R_m^i is non-increasing over time under each scenario and for each strain $i \in \mathcal{I}$, that is $R_m^i \leq R_{a(m)}^i$. This assumption reflects the decreasing production capability of the manufacturers as the flu season approaches. Furthermore, we assume that overall strain production quantity under each scenario is no more than κ, the total number of available chicken eggs.

Let $b(i, j)$ denote the cross–effectiveness between strains i and j, which we interpret as the ratio of immunity protection against strain j of a person who is vaccinated with strain i to that she would have gained if she was vaccinated with strain j. Note that b is not necessarily symmetric. The immune response to the flu shot is difficult to measure as it involves multiple cell types and depends on personal immunological history [122]. The Committee wants to ensure that there is a flu shot strain, which is antigenically close enough to each circulating strain. Therefore, we assume that the immunity protection provided by the flu shot against strain $j \in \mathcal{I}$ is determined by the vaccine strain that has the largest cross–effectiveness on j [10, 146].

Let binary variable $x_m^i = 1$ if strain $i \in \mathcal{I}$ is selected at node $m \in \mathcal{N}_{t<T}$, and $x_m^i = 0$ otherwise. Moreover, let w_n be the amount of vaccine produced at terminal node $n \in \mathcal{N}_T$. Producing more than the demand d_n is not profitable since the remaining vaccines at the end of the flu season cannot be used [74]. Therefore, d_n is a natural upper bound on $w_n \; \forall n \in \mathcal{N}_T$.

Let $\mathcal{H}(i, j) = \{k \in \mathcal{I} \mid b(k, j) > b(i, j)\}$, i.e. those strains that have higher cross–effectiveness on strain j than strain i does. We define variable $\delta_n^{ij} = w_n$ if none of the strains in $\mathcal{H}(i, j)$ are selected, and $\delta_n^{ij} = 0$ otherwise. We assume that no two strains have the same cross–effectiveness on a particular strain to ensure that only one flu shot strain is effective on a circulating strain.

Definition 3.1. *The "societal vaccination benefit" is the averted cost of infections by vaccinated people minus the cost of infections incurred by unvaccinated people. At terminal node* $n \in \mathcal{N}_T$, *it is given by* $F_n(\delta_n, w_n) = \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{I}} c_n e_n^j \delta_n^{ij} b(i, j) - q_n(d_n - w_n)$.

Note that optimizing $F_n(\delta_n, w_n)$ is equivalent to optimizing:

$$\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{I}} c_n e_n^j \delta_n^{ij} b(i,j) + q_n w_n = F_n(\delta_n, w_n) - q_n d_n, \tag{3.1}$$

as the second term on the right-hand side of (3.1) is constant and therefore dropped from the objective. Based on the notation defined above, the extensive form of the strain selection problem (SSP) is formulated as:

$$\text{[SSP]} \quad \max \sum_{n\in\mathcal{N}_T} p_n F_n(\delta_n, w_n) \tag{3.2}$$

$$\text{subject to} \sum_{m\in\mathscr{P}(n)\setminus\{n\}} x_m^i \le 1 \qquad \forall i\in\mathcal{I}, n\in\mathcal{N}_T, \tag{3.3}$$

$$\sum_{i\in\mathcal{I}}\sum_{m\in\mathscr{P}(n)\setminus\{n\}} R_m^i x_m^i \le \kappa \qquad \forall n\in\mathcal{N}_T, \tag{3.4}$$

$$w_n \le \sum_{i\in\mathcal{I}}\sum_{m\in\mathscr{P}(n)\setminus\{n\}} R_m^i x_m^i \qquad \forall n\in\mathcal{N}_T, \tag{3.5}$$

$$w_n \le \sum_{m\in\mathscr{P}(n)\setminus\{n\}} \min\{d_n, \beta_n^i R_m^i\} x_m^i + d_n\left[1 - \sum_{m\in\mathscr{P}(n)\setminus\{n\}} x_m^i\right] \qquad \forall i\in\mathcal{I}, n\in\mathcal{N}_T, \tag{3.6}$$

$$\delta_n^{ij} \le d_n\left[1 - \sum_{m\in\mathscr{P}(n)\setminus\{n\}} x_m^k\right] \qquad \forall i,j\in\mathcal{I}, k\in\mathcal{H}(i,j), n\in\mathcal{N}_T, \tag{3.7}$$

$$\delta_n^{ij} \le w_n, \quad \delta_n^{ij} \le d_n \sum_{m\in\mathscr{P}(n)\setminus\{n\}} x_m^i \qquad \forall i,j\in\mathcal{I}, n\in\mathcal{N}_T, \tag{3.8}$$

$$x_m^i \in \{0,1\}, \quad w_n, \delta_n^{ij} \ge 0 \qquad \forall i\in\mathcal{I}, m\in\mathcal{N}_{t<T}, n\in\mathcal{N}_T. \tag{3.9}$$

The objective function (3.2) maximizes the expected societal vaccination benefit. Initially, we assume a risk-neutral decision maker; hence, our objective function is based on the expectation of all possible outcomes. We incorporate risk sensitivity in Section 3.2. Constraints (3.3) allow a strain to be selected at most once under each scenario. Constraints (3.4) limit the total strain production to be no more than the number of eggs available. Constraints (3.5) ensure that $w_n = 0$ when no strain is selected. Note that if $w_n = 0$, the societal vaccination benefit is also zero. Constraints (3.6) state that the amount of vaccine produced

at terminal node n can not be more than the minimum production of all selected strains. Constraints (3.7) and (3.8) together ensure that only the vaccine strain with the highest cross–effectiveness provides immunity protection.

SSP may be viewed as a multi–stage stochastic capacity planning problem [63]. Strains correspond to facilities, and selecting a strain to include into the vaccine corresponds to selecting a facility to increase its capacity by a certain amount. Capacity planning problems typically have weak LP relaxations [169]. Hence, we reformulate SSP to tighten its LP relaxation. For all $i \in \mathcal{I}, n \in \mathcal{N}_T$ define variable $v_n^i = w_n$ if strain i is selected at node $n \in \mathcal{N}_T$, and $v_n^i = 0$ otherwise. Then, consider the following reformulation of SSP:

$$\text{[RSSP]} \quad \max \sum_{n \in \mathcal{N}_T} p_n F_n(\delta_n, w_n) \tag{3.10}$$

subject to (3.3), (3.4), (3.5), (3.7),

$$v_n^i \leq \sum_{m \in \mathscr{P}(n) \setminus \{n\}} \min\{d_n, \beta_n^i R_m^i\} x_m^i \qquad \forall i \in \mathcal{I}, n \in \mathcal{N}_T, \tag{3.11}$$

$$w_n \leq v_n^i + d_n \left[1 - \sum_{m \in \mathscr{P}(n) \setminus \{n\}} x_m^i\right] \qquad \forall i \in \mathcal{I}, n \in \mathcal{N}_T, \tag{3.12}$$

$$\delta_n^{ij} \leq v_n^i, \quad v_n^i \leq w_n \qquad \forall i, j \in \mathcal{I}, n \in \mathcal{N}_T, \tag{3.13}$$

$$x_m^i \in \{0,1\}, \quad w_n, v_n^i, \delta_n^{ij} \geq 0 \qquad \forall i, j \in \mathcal{I}, m \in \mathcal{N}_{t<T}, n \in \mathcal{N}_T. \tag{3.14}$$

Constraints (3.11) and (3.12) model constraints (3.6); and constraints (3.13) capture the relation enforced by constraints (3.8).

Proposition 3.1. *The optimal objective values of SSP and RSSP are equal.*

The proof of Proposition 3.1 is obvious and omitted. Note that RSSP has $|\mathcal{I}| \times |\mathcal{N}_T|$ more variables than SSP. However, its LP relaxation is stronger.

Proposition 3.2. *The optimal objective value of the LP relaxation of RSSP is no greater than that of SSP.*

Proof. Given a feasible solution $(\overline{x}, \overline{w}, \overline{\delta}, \overline{v})$ to the LP relaxation of RSSP, we construct a feasible solution (x, w, δ) to the LP relaxation of SSP with an objective value which is at least as large as that of RSSP. Consider a solution to SSP given by:

$$\begin{aligned} x_m^i &= \overline{x}_m^i, \\ w_n &= \min_{i\in\mathcal{I}} \left\{ \sum_{m\in\mathscr{P}(n)\setminus\{n\}} \min\{\beta_n^i R_m^i, d_n\}\overline{x}_m^i + d_n \left(1 - \sum_{m\in\mathscr{P}(n)\setminus\{n\}} \overline{x}_m^i\right)\right\}, \\ \delta_n^{ij} &= \overline{\delta}_n^{ij}. \end{aligned} \tag{3.15}$$

This solution is feasible to the LP relaxation of SSP. We show that $w_n \geq \overline{w}_n \ \forall n \in \mathcal{N}_T$.

$$\begin{aligned} w_n &= \min_{i\in\mathcal{I}} \left\{ \sum_{m\in\mathscr{P}(n)\setminus\{n\}} \min\{\beta_n^i R_m^i, d_n\}\overline{x}_m^i + d_n \left(1 - \sum_{m\in\mathscr{P}(n)\setminus\{n\}} \overline{x}_m^i\right)\right\}, \\ &\geq \min_{i\in\mathcal{I}} \left\{ \overline{v}_n^i + d_n \left(1 - \sum_{m\in\mathscr{P}(n)\setminus\{n\}} \overline{x}_m^i\right)\right\} \geq \overline{w}_n, \end{aligned} \tag{3.16}$$

where (3.16) follows from (3.11) and (3.12). As $\delta_n^{ij} = \overline{\delta}_n^{ij}$, the difference in the objective values of SSP and RSSP is given by $\sum_{n\in\mathcal{N}_T} p_n q_n (w_n - \overline{w}_n) \geq 0$. □

3.2 MEAN-RISK OBJECTIVES FOR RSSP

The vast majority of patients are risk averse over health outcomes [86], and public health officials may be as well. We extend our model to consider a risk-sensitive decision maker. A common approach to incorporate risk into stochastic programs is to consider a weighted mean-risk objective, where some deviation statistic is used as a measure of risk [162]. For example, the classical Markovitz [134] portfolio optimization model uses variance to measure risk; however, this criterion leads to non-convex objective functions in stochastic programs [175]. Another challenge related to stochastic programs with mean-risk objectives might be the loss of block-diagonal structure, which causes both computational and theoretical difficulties [84].

Risk measures that do not violate block-diagonal structure are discussed in [166] for two-stage stochastic programs and in [84] for multi-stage stochastic programs. Example

applications include portfolio optimization [49, 62], electricity networks optimization [113], and petroleum refinery planning [104]. We consider two risk measures from the literature: *the absolute semideviation* and *the expected shortage*, as we are concerned about the under performance of the influenza vaccine.

3.2.1 The Mean-Absolute Semideviation Objective

The absolute semideviation statistic measures the expected deviation below the mean. Given a coefficient $\gamma \in [0, 1]$ to trade off mean with risk, the mean-absolute semideviation objective for RSSP can be formulated as:

$$\text{MASD}(\gamma) = \left[(1-\gamma) \sum_{n \in \mathcal{N}_T} p_n F_n(\delta_n, w_n) - \gamma \sum_{n \in \mathcal{N}_T} p_n \max \left\{ \sum_{k \in \mathcal{N}_T} p_k F_k(\delta_k, w_k) - F_n(\delta_n, w_n), 0 \right\} \right].$$

To incorporate $\text{MASD}(\gamma)$ in RSSP, we define variable $\Gamma(n) \geq 0 \ \forall n \in \mathcal{N}_T$ that represents the absolute semideviation, and introduce:

$$\Gamma(n) \geq \sum_{k \in \mathcal{N}_T} p_k F_k(\delta_k, w_k) - F_n(\delta_n, w_n) \qquad n \in \mathcal{N}_T. \tag{3.17}$$

However, constraints (3.17) are not desirable as they create coupling among the terminal nodes. We generalize $\text{MASD}(\gamma)$ to reduce the coupling effect. For $1 \leq t \leq T$ define

$$\begin{aligned} \text{MASD}_t(\gamma) = \Bigg[& (1-\gamma) \sum_{n \in \mathcal{N}_T} p_n F_n(\delta_n, w_n) \\ & -\gamma \sum_{m \in \mathcal{N}_t} \sum_{n \in \mathcal{N}_T \cap \mathscr{T}(m)} p_n \max \left\{ \frac{1}{p_m} \sum_{k \in \mathcal{N}_T \cap \mathscr{T}(m)} p_k F_k(\delta_k, w_k) - F_n(\delta_n, w_n), 0 \right\} \Bigg], \end{aligned}$$

and replace (3.17) with

$$\Gamma_t(m, n) \geq \frac{1}{p_m} \sum_{k \in \mathcal{N}_T \cap \mathscr{T}(m)} p_k F_k(\delta_k, w_k) - F_n(\delta_n, w_n) \qquad m \in \mathcal{N}_t, n \in \mathcal{N}_T \cap \mathscr{T}(m). \tag{3.18}$$

It is easy to see that $\text{MASD}_1(\gamma) = \text{MASD}(\gamma)$, and $\text{MASD}_T(\gamma) = (1-\gamma)\ \text{MASD}(0)$. We interpret $\text{MASD}_t(\gamma)$ as the absolute semideviation starting from time stage t. Note that the coupling of terminal nodes induced by $\text{MASD}_t(\gamma)$ reduces as t increases.

3.2.2 The Mean-Expected Shortage Objective

The expected shortage statistic measures the expected deviations below a fixed target η. Given a coefficient $\gamma \in [0,1]$ to trade off mean with risk, the mean-expected shortage objective for RSSP can be formulated as:

$$\text{MES}(\gamma) = \left[(1-\gamma) \sum_{n \in \mathcal{N}_T} p_n F_n(\delta_n, w_n) - \gamma \sum_{n \in \mathcal{N}_T} p_n \max\left\{\eta - F_n(\delta_n, w_n), 0\right\} \right].$$

To incorporate MES(γ) in RSSP, we define variable $\Gamma(n) \geq 0 \ \forall n \in \mathcal{N}_T$, and introduce:

$$\Gamma(n) \geq \eta - F_n(\delta_n, w_n) \qquad\qquad n \in \mathcal{N}_T. \tag{3.19}$$

Note that MES(γ) does not create any coupling among the terminal nodes. Following Märkert and Schultz [133], we set η to the objective value of the wait-and-see solution [130] in our numerical experiments.

3.3 SOLUTION APPROACH

In RSSP, each strain selection decision $\{x_m^i \ \forall i \in \mathcal{I}, m \in \mathcal{N}_{T-1}\}$ corresponds to a unique solution of continuous variables that maximize the objective function (3.10). The general idea behind our solution approach is optimizing over the strain selection variables while using the continuous variables to calculate the objective value outside of the optimization algorithm.

3.3.1 Dantzig-Wolfe Reformulation of RSSP

We define binary variable $y^i_{\ell m} = 1$ if strain $i \in \mathcal{I}$ is selected at node $\ell \in \mathscr{P}(m)$ for node $m \in \mathcal{N}_{T-1}$, and $y^i_{\ell m} = 0$ otherwise. Let $\mathbf{y}_{\ell m} = (y^i_{\ell m})_{i\in\mathcal{I}}$ and $\mathbf{x}_\ell = (x^i_\ell)_{i\in\mathcal{I}}$. The strains selected for node $m \in \mathcal{N}_{T-1}$ at different time stages, i.e. $(\mathbf{y}_{\ell m})_{\ell\in\mathscr{P}(m)}$, correspond to a *vaccine design proposal* (VDP_m) for terminal nodes $n \in \mathscr{T}(m) \cap \mathcal{N}_T$. The set of all feasible VDP_m's for node $m \in \mathcal{N}_{T-1}$ is given by

$$\mathcal{Y}_m = \Big\{ y^i_{\ell m} \in \{0,1\},\ i \in \mathcal{I}, \ell \in \mathscr{P}(m) \ \Big|\ \sum_{\ell\in\mathscr{P}(m)} y^i_{\ell m} \le 1,\ i \in \mathcal{I},\ \sum_{i\in\mathcal{I}} \sum_{\ell\in\mathscr{P}(m)} R^i_\ell y^i_{\ell m} \le \kappa \Big\}.$$

Moreover, let $\mathcal{K}_m$ denote the index set of $\mathcal{Y}_m$, i.e. $\mathcal{Y}_m = \{(\hat{\mathbf{y}}_{\ell m})^k_{\ell\in\mathscr{P}(m)} \mid k \in \mathcal{K}_m\}$. Note that $\mathcal{K}_m$ is finite $\forall m \in \mathcal{N}_{T-1}$, and we can express any element of $\mathcal{Y}_m$ by

$$(\mathbf{y}_{\ell m})_{\ell\in\mathscr{P}(m)} = \sum_{k\in\mathcal{K}_m} \lambda^k_m (\hat{\mathbf{y}}_{\ell m})^k_{\ell\in\mathscr{P}(m)},\ \sum_{k\in\mathcal{K}_m} \lambda^k_m = 1,\ \lambda^k_m \in \{0,1\},\ k \in \mathcal{K}_m.$$

Note that each $\text{VDP}^k_m \equiv (\hat{\mathbf{y}}_{\ell m})^k_{\ell\in\mathscr{P}(m)}$ corresponds to an expected societal vaccination benefit $\sum_{n\in\mathscr{T}(m)\setminus\{m\}} p_n F_n(\hat{\delta}^k_n, \hat{w}^k_n)$ that maximizes the objective function (3.10), i.e. $\mathcal{K}_m$ indexes VDP_m's and societal vaccination benefits simultaneously. We use $(\mathbf{y}_{\ell m})_{\ell\in\mathscr{P}(m)}\ \forall m \in \mathcal{N}_{T-1}$ to obtain the Dantzig-Wolfe reformulation of RSSP

$$\text{[RSSP-DW]} \quad \max \sum_{m\in\mathcal{N}_{T-1}} \sum_{k\in\mathcal{K}_m} \left[\sum_{n\in\mathscr{T}(m)\setminus\{m\}} p_n F_n(\hat{\delta}^k_n, \hat{w}^k_n) \right] \lambda^k_m, \tag{3.20}$$

$$\text{subject to} \quad \sum_{k\in\mathcal{K}_m} \hat{\mathbf{y}}^k_{\ell m} \lambda^k_m = \mathbf{x}_\ell \qquad m \in \mathcal{N}_{T-1}, \ell \in \mathscr{P}(m), \tag{3.21}$$

$$\sum_{k\in\mathcal{K}_m} \lambda^k_m = 1 \qquad m \in \mathcal{N}_{T-1}, \tag{3.22}$$

$$\lambda^k_m \in \{0,1\},\ \mathbf{x}_\ell \ge 0 \qquad m \in \mathcal{N}_{T-1}, \ell \in \mathscr{P}(m), k \in \mathcal{K}_m. \tag{3.23}$$

Constraints (3.21) ensure that selected VDP_m's are nonanticipative. Convexity constraints (3.22) choose exactly one VDP_m for each node $m \in \mathcal{N}_{T-1}$. Note that RSSP-DW does not impose binary restrictions on $\mathbf{x}_\ell$ variables as they are satisfied for any binary vector λ.

The cardinality of $\mathcal{K}_m$ is huge, even for moderate-sized instances. We first create a restricted master problem (RSSP-RMP) in which each set $\mathcal{K}'_m$ represents a modest-sized

subset of $\mathcal{K}_m$. Given the optimal duals, $\hat{\pi}_{\ell m}$ and $\hat{\mu}_m$ that correspond to constraints (3.21) and (3.22) in the LP relaxation of the RSSP-RMP (RSSP-RMP-LP), we identify the column $\tilde{k} \in \mathcal{K}_m$ that has the most favorable reduced cost by solving the pricing subproblem

$$\text{[RSSP-SP(m)]} \quad \max \sum_{n \in \mathscr{T}(m) \cap \mathcal{N}_T} p_n F_n(\delta_n, w_n) - \sum_{i \in \mathcal{I}} \sum_{\ell \in \mathscr{P}(m)} \hat{\pi}^i_{\ell m} y^i_{\ell m} - \hat{\mu}_m, \tag{3.24}$$

$$\text{s.t.} \quad (\mathbf{y}_{\ell m})_{\ell \in \mathscr{P}(m)} \in \mathcal{Y}_m,$$

$$v^i_n \le \sum_{\ell \in \mathscr{P}(m)} \min\{d_n, \beta^i_n R^i_m\} y^i_{\ell m} \qquad i \in \mathcal{I}, n \in \mathscr{T}(m) \cap \mathcal{N}_T, \tag{3.25}$$

$$w_n \le v^i_n + d_n \left[1 - \sum_{\ell \in \mathscr{P}(m)} y^i_{\ell m}\right] \qquad i \in \mathcal{I}, n \in \mathscr{T}(m) \cap \mathcal{N}_T, \tag{3.26}$$

$$\delta^{ij}_n \le v^i_n, \quad v^i_n \le w_n \qquad i, j \in \mathcal{I}, n \in \mathscr{T}(m) \cap \mathcal{N}_T, \tag{3.27}$$

$$\delta^{ij}_n \le d_n \left[1 - \sum_{\ell \in \mathscr{P}(m)} y^k_{\ell m}\right] \qquad i, j \in \mathcal{I}, k \in \mathcal{H}(i,j), n \in \mathscr{T}(m) \cap \mathcal{N}_T, \tag{3.28}$$

$$\delta^{ij}_n, v^i_n, w_n \ge 0 \qquad i, j \in \mathcal{I}, n \in \mathscr{T}(m) \cap \mathcal{N}_T. \tag{3.29}$$

Note that RSSP-SP(m) is formulated for each node m in stage $T-1$, and $|\mathcal{N}_{T-1}| < |\mathcal{N}_T|$ unless each node in $|\mathcal{N}_{T-1}|$ has a single child. Any feasible solution to RSSP-SP(m) with a positive objective value $z_{RSSP-SP}(m) > 0$ leads to a new column for the RSSP-RMP, i.e. add a new element to $\mathcal{K}'_m$. If no such solution exists for any $m \in \mathcal{N}_{T-1}$, then we have solved the LP relaxation of RSSP-DW (RSSP-DW-LP) optimally. If the optimal solution to RSSP-DW-LP is integral, then we have solved RSSP-DW. If not, the algorithm branches. For further technical details of the branch-and-price method for multi-stage stochastic integer programs, we refer the reader to [129].

3.3.2 Branching Rule

Branching is more intricate in branch-and-price algorithms than that in traditional branch-and-bound algorithms, as it must prevent the regeneration of any columns previously fixed to zero while ensuring that the pricing algorithm used in the root is unchanged in the child nodes. In the literature, special branching rules have been developed for integer programming column generation [179].

Note that the only integer variables in RSSP-DW are λ_m^k's. However, there are numerous different values of λ_m^k variables that may result in the same strain selection decisions $\mathbf{x}_m \, \forall m \in \mathcal{N}_{t<T}$. Therefore, excessive enumeration may be required if we branch on λ_m^k variables. Moreover, the huge size of $\mathcal{K}_m$ exacerbates the situation.

Instead, we branch on the continuous $\mathbf{x}_\ell$ variables, which are to take binary values in the optimal solution. This approach has three main advantages over branching on λ_m^k variables. First, it enforces the nonanticipativity constraints (3.21). Second, it prevents regeneration of any columns previously fixed to zero as the branching constraints on $\mathbf{x}_\ell$ are enforced by modifying the bounds of $\mathbf{y}_{\ell m}$ variables in RSSP-SP(m). Finally, it has further implications as a strain can be selected at most once under each scenario:

$$x_\ell^i = 1 \Rightarrow x_{\ell'}^i = 0 \; \ell' \in (\mathscr{T}(\ell) \setminus \mathcal{N}_T) \cup (\mathscr{P}(\ell) \setminus \{\ell\}) \, .$$

We enforce the branching constraints by adjusting the bounds of $y_{\ell m}^i$ variables in the pricing problem. This ensures that the pricing algorithm used in the root is unchanged in the child nodes. Note that if x_m^i is branched to zero, then only the solutions such that $(\hat{y}_{\ell m}^i)^k = 0$ can be selected from $\mathcal{K}'_m$. Hence we can reduce the size of the RSSP-RMP at each node of the branch-and-bound tree by eliminating all columns whose components do not satisfy the new branching constraints in the successor nodes.

3.3.3 Solving the Pricing Subproblems

We solve RSSP-SP(m) using dynamic programming, which has two main advantages over using an off-the-shelf MIP solver, as was verified by preliminary computations. First, it searches over the $\mathbf{y}_{\ell m}$ variables and calculates the corresponding values of the continuous variables in RSSP-SP(m) exogenously, which is usually much faster than optimizing over the whole mixed-binary variable space, particularly when there are few time stages and many scenarios. Second, it enables us to consider all feasible columns generated throughout the dynamic programming iterations, which may improve convergence of the column generation.

At any node g of the branch-and-price tree, let $\mathcal{B}_m^+(g)$ be the set of all $(i, \ell) \in (\mathcal{I} \times \mathscr{P}(m))$ pairs such that $y_{\ell m}^i$ has been fixed to one by prior branching decisions. Likewise, let $\mathcal{B}_m^-(g)$ be

the set of all $(i, \ell) \in (\mathcal{I} \times \mathscr{P}(m))$ pairs such that $y^i_{\ell m}$ has been fixed to zero. In our dynamic programming formulation of SP(m), there are $|\mathcal{I}| + 1$ stages. Moreover, state s_t at stage t is defined by an $|\mathcal{I}|$-dimensional vector whose i^{th} component $s_t(i)$ is equal to $\ell \in \mathscr{P}(m)$ if strain i is selected at node ℓ and 0 otherwise. Let $\mathbf{S}$ be the entire state space. Note that the initial state s_1 is uniquely defined by $\mathcal{B}^+_m(g)$ and $\mathcal{B}^-_m(g)$.

We denote the remaining production capacity in state s_t by $\kappa(s_t)$. Then, at stage $t \leq |\mathcal{I}|$, the action space $\mathcal{A}(s_t)$ is empty if for $i = t$, $s_t(i) \neq 0$, i.e. strain i was selected before. Otherwise, it consists of nodes $\ell \in \mathscr{P}(m)$ such that $R^i_\ell \leq \kappa(s_t)$ and $(i, \ell) \notin \mathcal{B}^-_m(g) \cup \mathcal{B}^+_m(g)$. Given decision $a \in \mathcal{A}(s_t)$ in state s_t, we denote the state at the next stage by $(s_t + a)$. Note that this definition of $\mathcal{A}(s_t)$ guarantees that starting from an initial state s_1 that corresponds to a feasible solution to SP(m), all subsequently visited states from $t = 1$ to $|\mathcal{I}| + 1$ also correspond to feasible solutions. Let the objective value (10a) of the feasible solution that corresponds to $s_{|\mathcal{I}|+1}$ be equal to $J_m(s_{|\mathcal{I}|+1})$. Then for $t = |\mathcal{I}|, \ldots, 1$, the optimal benefit-to-go function is defined by:

$$J_m(s_t) = \max_{a \in \mathcal{A}(s_t)} \Big[J_m(s_t + a) \Big] \qquad \forall s_t \in \mathbf{S}. \tag{3.30}$$

It is easy to see that the dynamic programming algorithm defined above terminates finitely after $O(T|\mathcal{I}|)$ iterations, and the optimal objective value of SP(m) is equal to $J_m(s_1)$.

3.3.4 Solving the Mean-Risk Models

We solve the mean-risk variants of RSSP by modifying the pricing subproblem RSSP-SP(m). To the best of our knowledge, ours is the first branch-and-price algorithm proposed for a multi-stage mean-risk model. For the mean-absolute deviation objective we consider $\text{MASD}_{T-1}(\gamma)$, as it creates the least coupling among the terminal nodes. We solve the following pricing subproblem $\forall m \in \mathcal{N}_{T-1}$.

$$\max \text{MASD}_{T-1}(\gamma) - \sum_{i \in \mathcal{I}} \sum_{\ell \in \mathscr{P}(m)} \hat{\pi}^i_{\ell m} y^i_{\ell m} - \hat{\mu}_m, \tag{3.31}$$

$$\text{subject to} \quad (\mathbf{y}_{\ell m})_{\ell \in \mathscr{P}(m)} \in \mathcal{Y}_m,$$

$$y^i_{\ell m} \in \{0, 1\}, \delta^{ij}_n, v^i_n, w_n, \Gamma(n) \geq 0 \, \forall i, j \in \mathcal{I}, \ell \in \mathscr{P}(m).$$

For the mean-expected shortage objective, we use MES(γ) in the objective function (3.31). Either versions of the subproblem can easily be incorporated in our branch-and-price method.

3.4 PARAMETRIZATION

Our parameter values are based on the 2008-09 flu season in the United States. Nine candidate strains were considered in the Committee's meeting in February 2008. As seen in Table 1.1, four of them were A/H1N1 subtypes, two of them were A/H3N2 subtypes, and three of them were influenza B viruses. We calibrate our model using values from the literature and/or from publicly available data. We use Monte Carlo sampling to generate point estimations of these parameters under each scenario.

Vaccine demand and the attack rate. As seen in Figure 1.2, there is strong correlation between the amounts of the flu shot produced and distributed. Therefore, estimating the demand based on historical data is not reasonable. According to CDC recommendations, influenza vaccination target population size is 281.1 million people in the United States in 2006 [30]. Cho [45] uses a base demand of 103 million doses (36.6% of the target) as individuals do not internalize their indirect costs of infection.

The attack rate is the proportion of individuals exposed to an infectious agent who become clinically ill. If the influenza attack rate is expected to be higher than in a typical season, the expected cost of infection incurred by unvaccinated individuals (q_n) and the flu shot demand (d_n) increases. We use an aggregate demand that is uniformly distributed between $30-40\%$, $40-50\%$ and $50-60\%$ of the CDC's target population size when the influenza attack rate is low, moderate and high, respectively.

Production yield ratios. We are unaware of any publicly available data from the FDA or flu shot manufacturers about the production yields of strains. Moreover, we are unaware of any such study in the literature. As seen in Table 1.1, the Committee classifies the candidate flu shot strains as low, moderate and high yield based on the availability and efficiency of the high-growth seed strains. Cho [45] uses a baseline value of 0.9 for the production yield ratio. We assume that production yield ratios have a truncated-normal distribution between 0–1,

and set the low yield ratio mean to 0.7, moderate yield ratio mean to 0.8, and high yield ratio mean to 0.9. The production yields are more predictable if the current flu shot composition is retained. Chick et al. [44] uses a baseline value of 0.2 for the standard deviation of the production yield ratios. Cho [45] sets this value to 0.1 for the current vaccine strains and to 0.4 for the others. Conservatively, we set the standard deviation of the production yield ratios to 0.2 for the current vaccine strains, and to 0.4 for the others.

Egg supply. Producing $281.1 \times 0.6 \cong 170$ million doses of trivalent influenza vaccine to satisfy the maximum possible demand in our model, would require $(170/0.8) \times 3 \cong 640$ million fertilized eggs [89] under the medium yield ratio. This number would presumably increase if more than three strains are included in the flu shot. In 2004, the Department of Health and Human Services (HHS) contracted with Sanofi Aventis Inc. to ensure that there are enough eggs to manufacture the flu shot in the event of a pandemic or shortage [87]. Therefore, the egg supply is known before the strain selections. We assume a constant egg supply of $\kappa = 750$ million eggs under all scenarios.

Strain production quantities. There are no published data available to differentiate between the number of chicken eggs used for the production of different strains. Therefore, we omit the strain index i from R_m^i. All three strains for the 2008-09 flu season were selected in the first meeting in February 2008 [67]. The manufacturers projected to produce as many as 146 million doses of flu shot for the United States market [90] for the 2008-09 flu season. Therefore, R_1, the strain production quantity at the root node of the scenario tree should be $\sim 146/0.8 = 182.5$ million under the moderate production yield ratio. Manufacturers might not be producing at their highest capacity, as they are profit maximizers. We assume that R_1 has a truncated-normal distribution between $180 - 200$ million eggs with a mean of 190 million eggs and a standard deviation of 5 million eggs.

Based on a 24-week continuous flu shot production [45], the average production rate (i.e. egg processing rate) would be $\approx 190/24 = 7.91$ million eggs per week. However, the time required for the preparation of the high-growth seed strains and testing prior to lot release is about the same regardless of the timing of strain selection decisions. Therefore, deferring the decisions shortens the time allocated for monovalent strain production. Around one month is required for the preparation of high-growth seed strains, and another one month is required

for filling, packaging and testing prior to lot release [69]. Hence, to process 190 million eggs, weekly production rate should be as high as $\approx 190/16 = 11.87$ million eggs on the basis of a 16-week continuous monovalent strain production campaign.

We assume that the production rate is uniformly distributed between 7.91 and 11.87 million eggs per week across different scenarios. Let r_m be the realization of this rate at node $m \in \mathcal{N}_{t<T}$, and Δt be the number of weeks between two consecutive Committee meetings. Then, we estimate the available strain production capability left at node m as $R_m = R_{a(m)} - \Delta t \times r_m$ million eggs.

Cost of infection and vaccine shortages. We assume that the expected cost of flu infection c_n, which includes the expected health care cost, work loss cost and lost earnings, is \$41 per person under all scenarios [183]. Shortage cost q_n represents the expected cost of infection incurred by an unvaccinated individual. According to the CDC [39], $5-20\%$ of the population gets the flu in the United States in a typical flu season. A high influenza attack rate would increase the expected cost of infection incurred by unvaccinated individuals. We let q_n vary uniformly between $\$41 \times 0.05 = \$2.05$ and $\$41 \times 0.1 = \$4.1$ per person when the influenza attack rate is low, between $\$41 \times 0.1 = \$4.1$ and $\$41 \times 0.2 = \$8.2$ per person when the influenza attack rate is moderate, and between $\$41 \times 0.2 = \$8.2$ and $\$41 \times 0.4 = \$16.4$ per person when the influenza attack rate is high.

Prevalence of strains. The *relative prevalence* of a strain is the ratio of the number of people who suffered from that particular strain to the overall infected population size. We consider three relative prevalence levels as low (0-20%), medium (20-50%) and high (>50%). Note that the uncertainty about the relative prevalence of strains reveals gradually according to a trend, e.g. low to medium, high to high. A straightforward point estimate that is equal to the average of past observations might be misleading. We develop an ordinal logistic regression model, which estimates one stage transition probabilities among the low, medium and high relative prevalence levels of a strain based on its current relative prevalence level (χ_1) and a moving average of the past prevalence observations (χ_2).

Let π_h, π_m, and π_ℓ denote the transition probabilities to high, medium, and low classes. As the outcome variable has three levels, the ordinal logistic regression model has two equations with two different intercepts, i.e. $Y_i = \zeta_i + \beta_1\chi_1 + \beta_2\chi_2$ for $i = 1, 2$, and the transition probability equations are given by

$$\pi_\ell = \frac{\exp(Y_1)}{1+\exp(Y_1)},\ \pi_m = \frac{\exp(Y_2)}{1+\exp(Y_2)} - \frac{\exp(Y_1)}{1+\exp(Y_1)},\ \pi_h = 1 - \frac{\exp(Y_2)}{1+\exp(Y_2)}$$

.

Every flu season is unique due to antigenic changes of the flu viruses. We assume that historical trends in the prevalence of A/H1N1, A/H3N2 and B strain categories would represent the prevalence of specific strains in those categories. Our estimations are based on the past weekly surveillance reports of the CDC, which date back to 1997 and contain the prevalence of circulating flu strains between week 40 of one calender year and week 20 of the subsequent calender year. We summarize the estimated parameters of our logistic regression models for each three strain categories in Table 3.1.

We initialize the relative prevalence level of each strain to medium. Then using Monte Carlo sampling we vary each strain's relative prevalence level independently for 32 weeks, i.e. from March to October. Once relative prevalence levels of strains in the flu season have been determined, we make point estimations from a uniform distribution between $0-20\%$ for the low level strains, $20-50\%$ for the medium level strains, and $50-80\%$ for the high level strains. Note that prevalence of all candidate strains can be no more than 100%. Therefore, we normalize this sum when necessary.

As mentioned before, $5-20\%$ of the population gets the flu in the United States in a typical flu season [39]. We multiply the point estimations of the relative prevalence of strains by a coefficient which is uniformly distributed between 0.05 and 0.1 when the influenza attack rate is low, between 0.1 and 0.2 when the influenza attack rate is moderate and between 0.2 and 0.5 when the influenza attack rate is high.

Cross-effectiveness. The HI assays, which are used by the Committee to evaluate cross–effectiveness among the strains, provide only rank order information. In the literature, the most widely used methods to quantify cross–effectiveness include comparing genetic

sequences based on Hamming distance [123], constructing phylogenetic trees [171], and analyzing epitopes [141]. Following Wu et al. [191], we use the *shape space* model of Perelson and Oster [150], which represents antigens and antibodies by unique vectors in Euclidean space (the shape space). The coordinates of vectors are assumed to characterize binding properties.

Lapades and Farber [117] uses an ordinal multidimensional scaling algorithm [111, 112] to model the dimension of the shape space, and the relative coordinates of strains from an HI assay panel. If there are M antigens and N antibodies in the panel, then there are MN experimental values, E_{ij}, which represent the cross–effectiveness between antigen i and antibody j. These values are assumed to be monotonically related to the distance D_{ij} in the shape space. The E_{ij} values are sorted in descending order and indexed with a number α so that $E_{\alpha+1} \leq E_\alpha$ for $\alpha = 1, \ldots, MN-1$. The cross–effectiveness of an antigen-antibody pair is inversely proportional to their antigenic distance. The ordinal multidimensional scaling algorithm finds shape space coordinates of antigens and antibodies so that the orderings of D_{ij} and E_{ij} are reversed, i.e. $D_{\alpha+1} \geq D_\alpha$ for $\alpha = 1, \ldots, MN-1$. We formulate a nonlinear program that sorts the D_α values in the desired order according to the experimental values:

$$\max \quad \sum_{\alpha=1}^{MN} \frac{1}{2} \left[1 + \tanh(D_{\alpha+1} - D_\alpha)\right] \tag{3.32a}$$

$$D_\alpha \leq D_{\alpha+1} \qquad \alpha = 1, \ldots, MN-1. \tag{3.32b}$$

Note that each term of the summation in objective function (3.32a) tends to one as $D_{\alpha+1} - D_\alpha$ increases. Thus, the global optimal objective value is equal to MN. We use objective function (3.32a) to eliminate trivial solutions that assign zero to all shape space coordinates. However, different objective functions can be used since the exact algebraic form is not critical for the shape space model [117]. We solve problem (3.32) using the off-the-shelf nonlinear solver KNITRO [106] with different starting points. We do not search for the global optimal solution as any feasible solution that preserves the corresponding rank ordering is sufficient.

We assume that a strain is a hundred percent effective on itself, i.e. $b(i,i) = 1, \ \forall i \in \mathcal{I}$. Given the shape space distances, we calculate the cross-effectiveness between antibody i and

antigen j by $b(i,j) = 1 - D_{ij}/D_{\max}\ \forall i \neq j$, where $D_{\max}$ be the largest shape space distance. We summarize our estimations for the cross-effectiveness among candidate vaccine strains in Table 3.2. For example, we interpret the third entry in the first row that a person who is vaccinated with A/Brisbane/59/07 strain against A/New Caledonia/20/99 antigen will gain 0.278 of the immunity protection that she would have gained if she was vaccinated with A/New Caledonia/20/99 strain.

3.5 NUMERICAL EXPERIMENTS

3.5.1 Implementation

We implement our branch-and-price algorithm using BCP, a framework for branch, cut, and price algorithms [155]. BCP handles all bookkeeping related to search tree management, column and cut generation, branching, and so on. We initialize the restricted master problem using a greedy heuristic. Several problem specific adjustments are required to make the branch-and-price algorithm work efficiently. In particular, when processing a node, we do not always solve the RSSP-DW-LP optimally. If the convergence is slow, or if $z_{LP}^{RSSP-RMP}$, the objective value of the the RSSP-RMP-LP, is greater than the current local upper bound, then we stop generating columns and branch. Moreover, we apply a dual stabilization technique [61] to restrict the fluctuations of dual variables and improve convergence.

Given a solution to the RSSP-DW-LP, we branch on the most fractional x_m^i variable. When a node is either infeasible or fathomed, we select the node with the best upper bound. Otherwise, when we branch, we dive with the hope of finding a new lower bound. At each node of the branch-and-price algorithm, we solve the RSSP-RMP-LP using CLP, an open-source LP solver [72].

3.5.2 Performance of the Proposed Solution Technique

We evaluate the performance of the proposed solution technique using 12 SSP instances whose characteristics are provided in Table 3.3. As seen in Figure 3.2, the strain selection

time horizon is equal to 4 weeks in 3-, 4- and 6-stage instances, and 4.5 weeks in 5-stage instances. We set $\Delta t = 1.5$ weeks for 5-stage instances as a 4/3-week time interval may not be practical.

Table 3.4 reports the performance of our branch-and-price algorithm against CPLEX 11.0 (2010) with default settings. All computational experiments are conducted on an Intel Xeon PC with 3GHz CPU and 3GB of RAM using ten-hour time limit. CPLEX 11.0 (2010) cannot solve the extensive form of our instances within the time limit, except the smallest ones with three stages, while the proposed branch-and-price algorithm solves all instances in less than two hours. Note that RSSP-LP is much tighter than the SSP-LP for all test instances. Moreover, MP-LP, the LP relaxation of the MP at the root node of the branch-and-price tree, is often integer feasible.

3.6 EVALUATING POLICY ISSUES

3.6.1 Value of Integrating Timing and Composition Decisions

In this section, we evaluate the value of relaxing the two common assumptions in the literature as mentioned in Section 1.4: (*i*) composition decisions are independent among different strain categories [45, 109, 191], (*ii*) the most prevalent circulating strain is the only alternative for the current vaccine strain [45, 109].

Considering each strain category simultaneously. To quantify the benefits of considering all three strain categories simultaneously, we relax constraint (3.12), which limits the flu shot production to be the minimum of all selected strain productions. Moreover, we define a production variable for each strain category. After solving the relaxed model, we reimpose (3.12) belatedly. As seen in the "Separate Categories" column of Table 3.5, considering each strain category separately does not reduce the optimal objective function much. This result is mainly due to our assumption that there is no cross-effectiveness among strains in different categories (see Table 3.2). However, the resulting solutions are still not optimal, as the overall production quantity of the multi-valent flu shot is not considered.

Considering extra candidate strains in addition to the most prevalent ones. To quantify the benefits of considering extra candidate strains beside the most prevalent ones, we solve our instances by considering only the strains that have the highest current expected prevalence in their category. As seen in the "Highest Exp. Prevalence" column of Table 3.5, considering only strains that have the highest current expected prevalence in their category may result in poor solutions, e.g. the H1 instance. This indicates that selecting a less prevalent strain might be favorable if it has a better production yield and/or cross–effectiveness.

3.6.2 Policy Implications

There are various opportunities to ameliorate flu epidemics, e.g. improving surveillance, incorporating more strains into the flu shot, and reducing the production yield uncertainties. Our model can quantify the societal benefit of these opportunities and valuable empirical insights can be gained by numerical experiments.

More frequent decision epochs. Currently, the Committee makes at most two meetings with a four-week interval; one in February and the other is in March [69]. The strain selection decisions should not be deferred beyond March due to the lengthy manufacturing process. However, the Committee could meet more often to enable earlier commitments. We consider different durations between consecutive Committee meetings (Δt) for the instances that have different number of time stages. As seen in Table 3.5, the optimal objective value of RSSP often increases as the number stages increases. This result indicates that more frequent Committee meetings could provide higher flexibility when selecting the strains with respect to the uncertainties.

Influenza attack rate. If the influenza attack rate is expected to be higher than that in a typical season, the expected cost of infection incurred by unvaccinated individuals (q_n) and the flu shot demand (d_n) increases. In Table 3.3, there are four instances under each attack rate: low, medium and high. An immediate and intuitive observation from Table 3.5 is that the higher the attack rate, the greater the benefit gained from the flu shot.

The number of strains in the flu shot. National authorities could adjust the number and/or dosage of strains in the flu shot. For instance, the FDA recently considered incorporating a fourth strain. The following quote is from a flu shot manufacturer [68]:
"We believe that moving to a routine four strain recommendation for annual flu vaccination, two A strains, two B strains, while adding some manufacturing complexity, has substantial benefits. We agree, production capacity across the industry is adequate to support this initiative, and inclusion of both B lineages will address the recurring issue."

Our model does not restrict the number of strains in the flu shot. More strains would provide higher immune response but the flu shot would be more expensive and more difficult to manufacture. To estimate the benefits of flexibility in strain selection, we solve our instances by enforcing the selection of a single strain for each category. Formally, let $\mathcal{I}_1$, $\mathcal{I}_2$ and $\mathcal{I}_3$ be the index set of strains in the A/H1N1, A/H3N2 and B categories, respectively. Then, for each strain category $k = 1, 2, 3$, we replace constraint (3.3) with

$$\sum_{m \in \mathscr{P}(n) \setminus \{n\}} x_m^i = 1 \; i \in \mathcal{I}_k, n \in \mathcal{N}_T.$$

As seen in the "Single Strain per Category" column of Table 3.5, enforcing the selection of a single strain for each category does not result in substantially worse solutions. This result is due to our assumption that only the vaccine strain that has the highest cross-effectiveness provides immunity protection against a particular virus strain. Note that when there are two strains selected from a category, only one of them is effective on each circulating strain from that category. Nevertheless, selecting more than one strain from a category might still be favorable because either of these strains might have higher cross-effectiveness on a particular virus strain.

Hedging against production yield uncertainties. Unanticipated problems in growing a single vaccine strain might result in reductions of the overall flu shot supply. Currently, the Committee utilizes a fixed and somewhat vague classification of production yield ratios (see Table 1.1). In contrast, our model allows production yield ratios to vary across scenarios. To quantify the impact of this modeling flexibility, we solve our instances by assuming that the production yield ratio of each strain is equal to its mean over all scenarios. We then substitute the optimal solutions of these modified instances into the original model. As seen

in the "Expected Yield" column of Table 3.5, ignoring production yield uncertainties might result in very poor solutions under all attack rates.

Greater strain production quantities. Currently, egg-based production of flu strains takes about six months [74]. As a remedy, researchers have proposed development of cell-culture-based vaccine production technology [89, 103]. If feasible, cell-culture technologies would offer distinct advantages over egg-based manufacturing methods, e.g. reduced contamination risk, shorter production start-up times, greater production rates, to name a few [156], which would enable greater strain production quantities.

Greater strain production quantities would provide flexibility in strain selections (i.e. selecting prevalent strains which have low production yield ratios, or including additional strains into the flu shot) and allow a longer surveillance period (i.e. waiting for more data before selecting the flu shot strains). We perform a sensitivity analysis by solving our instances with 10%, 20%, and 30% increased strain production quantities and egg supply. Figure 3.3 reports the results of this experiment. Note that a 30% increase in the strain production quantities leads to more than 20% increase in the societal benefit. We therefore concur with Deo and Corbett [56] that it may be in society's interest to subsidize research into such manufacturing techniques.

3.6.3 The Mean-Risk Efficient Frontier

We solve the mean-risk variants of RSSP repeatedly for different values of risk coefficient γ to approximate the mean-risk efficient frontier. It is well known that this approach may not identify all Pareto efficient solutions [173]. However, the approximation quality can be enhanced by solving more instances with different risk coefficients γ.

We consider three four-stage RSSP instances under the low, medium, and high attack rates (i.e. L2, M2, H2 instances). We solve each instance using the mean-absolute semideviation and the mean-expected shortage objectives with 18 equidistant values of γ from 0.05 to 0.9. Figures 3.4, 3.5, and 3.6 depict the mean-risk efficient frontiers under different attack rates. We observe that the expected vaccination benefit becomes more sensitive to risk as the attack rate gets higher.

In Figures 3.5 and 3.6, the Pareto efficient mean-risk points are clustered in a few groups. Two phenomenons may cause this result. Consider three different values of the risk coefficient: $0 \leq \gamma' < \gamma_0 < \gamma'' \leq 1$. When γ increases from γ_0 to γ'', the optimal mean-risk pair, or the optimal flu shot composition, may stay the same if the expected cost of updating the solution is more than the benefit of corresponding risk decrease. Likewise, when γ decreases from γ_0 to γ', the optimal mean-risk pair may stay the same if the expected benefit of updating the current solution is less than the cost of corresponding risk increase. We note that the numerical experiments in [1, 2] on stochastic linear programs showed similar behavior.

3.7 CONCLUSIONS

We find that more frequent Committee meetings can provide up to 10% gains in the annual societal benefit of the flu shot. Incorporating more than three strains in the flu shot can increase the annual societal benefit by more than $80 million, particularly under more severe flu seasons. Complying with the current practice we recommend the use of multiple candidate strains per category, as a strain, which is not necessarily the most prevalent one, might be favorable based on its manufacturability and cross-effectiveness properties. However, different strain categories should be considered simultaneously when choosing the composition of the influenza vaccine, because unanticipated difficulties growing a strain might result in reductions of the overall vaccine supply. The Committee utilizes a fixed and somewhat vague classification of production yield ratios (see Table 1.1). We find that ignoring production yield uncertainties might result in very poor vaccine designs. Finally, we find that enhanced manufacturing techniques have substantial benefits, so that it may be in society's interest to subsidize research into such manufacturing techniques.

We model the perspective of the Committee, rather than that of the manufacturers. The Committee's objective is to maximize the societal vaccination benefit, whereas the manufacturers are profit maximizers. However, these two objectives are closely interconnected. The manufacturers' profits are affected by the strain selection decisions of the Committee because not all strains grow well in chicken eggs. Moreover, if the selected strains match well with

the virus strains that emerge during the epidemic season, there would be higher demand for the flu shot. The societal vaccination benefit depends on the coverage rates and timely availability of the vaccine. In Chapter 4, we propose a model that incorporates both the Committee's and the manufacturers' perspectives. That model decides on the production amounts dynamically in response to selected strains.

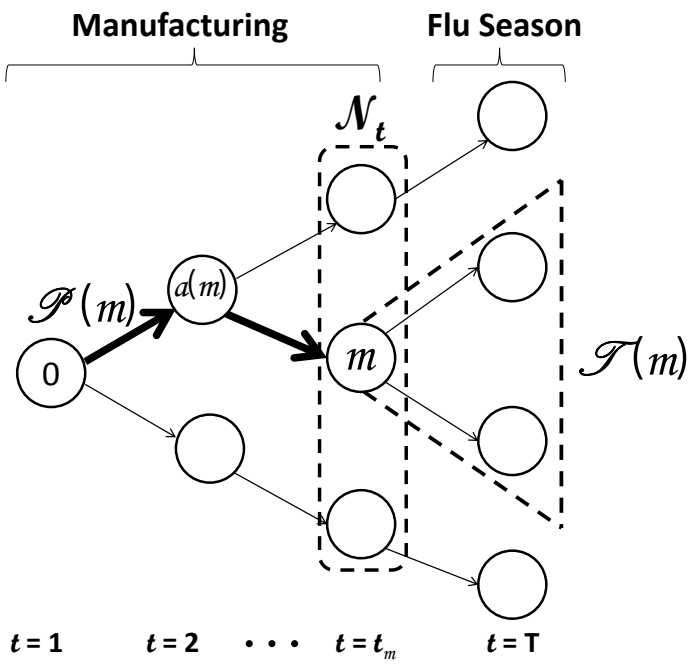


Figure 3.1: Each node m, except the root, is joined to its ancestor $a(m)$. The set of nodes at time stage t is represented by $\mathcal{N}_t$. The probability of node m is p_m where $\sum_{m \in \mathcal{N}_t} p_m = 1$ for all $t \leq T$. Moreover, $\mathscr{P}(m)$ is the unique path linking the root node to m and $\mathscr{T}(m)$ is the subtree rooted at m. Paths between the root node and the terminal nodes correspond to individual scenarios. Thus, nodes at level t represent possible scenario realizations at time stage t, and terminal nodes are also known as scenarios.

Table 3.1: Estimated parameters of the ordinal logistic regression model.

Parameter	A/H1N1	A/H3N2	B
ζ_1	4.0995	5.0960	4.1833
ζ_2	8.1723	10.0081	9.4719
β_1	-0.1606	-0.1733	-0.2302
β_2	-1.4036	-1.7736	-1.6331

Table 3.2: Cross–effectiveness among the candidate strains. Blank entries indicate no cross-effectiveness.

	Strain Name	(S1)	(S2)	(S3)	(S4)	(S5)	(S6)	(S7)	(S8)	(S9)
(S1)	A/New Caledonia/20/99	1.000	0.333	0.278	0.266					
(S2)	A/Solomon Islands/3/06	0.341	1.000	0.498	0.261					
(S3)	A/Brisbane/59/07	0.232	0.329	1.000	0.614					
(S4)	A/South Dakota/6/07	0.227	0.325	0.661	1.000					
(S5)	A/Wisconsin/67/2005					1.000	0.221			
(S6)	A/Brisbane/10/2007					0.265	1.000			
(S7)	B/Malaysia/2506/2004							1.000	0.097	0.054
(S8)	B/Florida/04/2006							0.339	1.000	0.603
(S9)	B/Brisbane/03/2007							0.334	0.462	1.000

Table 3.3: Sizes of the test instances. The attack rate is the proportion of individuals exposed to an infectious agent who becomes clinically ill. The scenarios used under different attack rates are the same except the demand (d_n), shortage cost (q_n) and the prevalence (e_n^i) parameters as explained in Section 3.4. The number of weeks between two consecutive Committee meetings is denoted by Δt.

Attack		Δt	Scenario tree size			Extensive form size		
rate	Name	(weeks)	$\|T\|$	$\|\mathcal{N}_T\|$	$\|\mathcal{N}\|$	Cont. Var.	Bin. Var.	Constraints
Low	L1	4	3	100	111	1000	99	84800
	L2	2	4	512	585	5120	657	434176
	L3	1.5	5	243	341	2430	882	206064
	L4	1	6	32	63	320	279	27136
Moderate	M1	4	3	100	111	1000	99	84800
	M2	2	4	512	585	5120	657	434176
	M3	1.5	5	243	341	2430	882	206064
	M4	1	6	32	63	320	279	27136
High	H1	4	3	100	111	1000	99	84800
	H2	2	4	512	585	5120	657	434176
	H3	1.5	5	243	341	2430	882	206064
	H4	1	6	32	63	320	279	27136

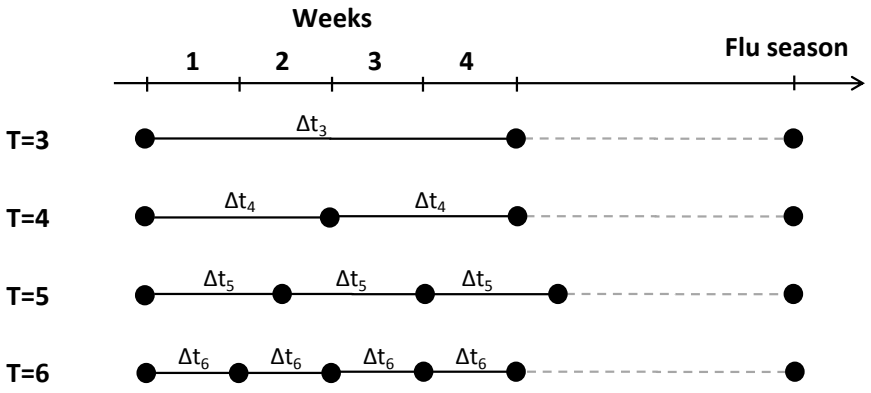


Figure 3.2: Time stages of 3-, 4-, 5- and 6-stage models. Composition decisions are made in the first $T-1$ stages, and the uncertainties about the characteristics of the flu season are realized in the last stage.

Table 3.4: Performance of the proposed solution technique. RRSP-DW-LP is the LP relaxation of the RSSP-DW at the root node of the branch-and-price tree. % gap is calculated relative to the optimal solution. A '-' indicates that the 10-hour time limit was exceeded.

	SSP-LP	RSSP-LP	RSSP-DW-LP	CPLEX	Branch-and-price
	% gap	% gap	% gap	sec.	sec.
L1	13.21	7.59	< 0.01	323.67	5.53
L2	12.74	8.01	0.00	-	853.93
L3	13.69	8.70	0.00	-	5121.71
L4	9.38	6.48	< 0.01	-	3848.07
M1	19.63	10.63	0.00	251.05	3.84
M2	21.60	12.07	0.00	-	682.88
M3	22.38	13.39	0.00	-	3561.06
M4	20.60	12.94	0.00	-	2533.88
H1	38.76	20.75	0.00	220.27	3.28
H2	40.56	21.70	0.00	-	439.59
H3	41.02	23.45	0.00	-	2232.56
H4	33.26	19.91	< 0.01	-	3211.83

Table 3.5: Experiments for policy issues. The columns expressing a percentage are relative to the optimal solution.

	Optimal val. RSSP ($M)	Separate Categories (%)	Highest Exp. Prevalence (%)	Single Strain per Category (%)	Expected Yield (%)
L1	432.02	99.83	88.03	98.58	63.82
L2	444.71	99.49	92.66	97.50	71.71
L3	447.96	99.63	90.37	95.74	65.57
L4	474.95	99.44	91.13	94.41	65.99
M1	1022.92	99.89	92.24	99.77	59.73
M2	1020.45	99.76	95.19	99.61	59.31
M3	1033.36	99.42	95.59	98.30	59.96
M4	1074.32	97.86	87.74	96.11	55.61
H1	2343.53	99.97	85.95	99.97	56.08
H2	2309.87	99.88	89.63	99.88	58.11
H3	2319.21	99.30	95.54	99.00	57.67
H4	2530.17	94.03	93.92	96.72	58.89

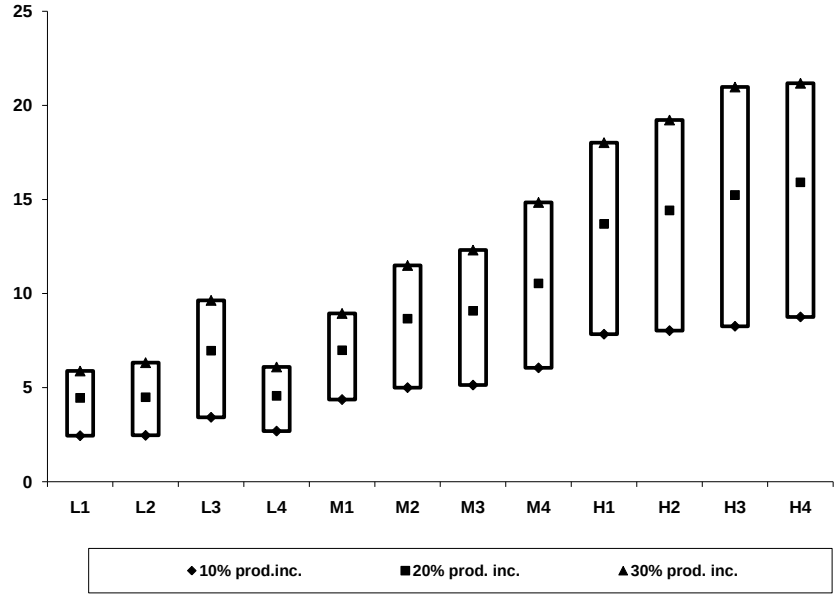
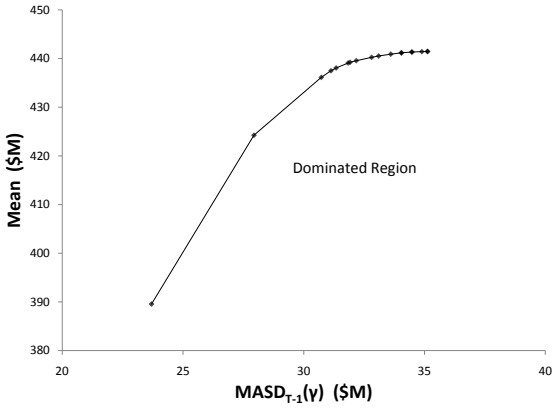


Figure 3.3: % improvement in the current optimal objective function value due to increasing the strain production capacity.



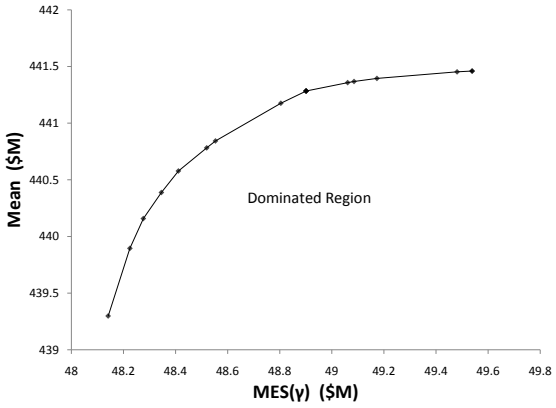
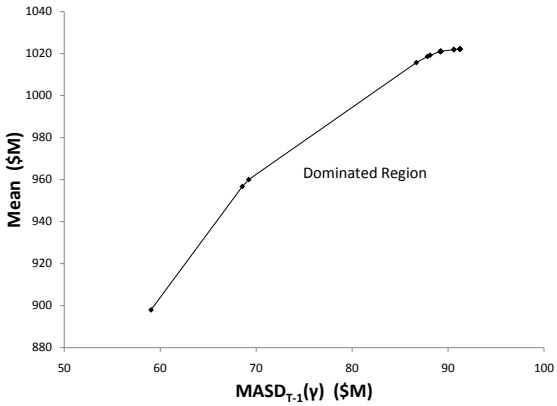


Figure 3.4: The absolute semideviation and the expected shortage mean-risk approximate efficient frontiers under the low attack rate.



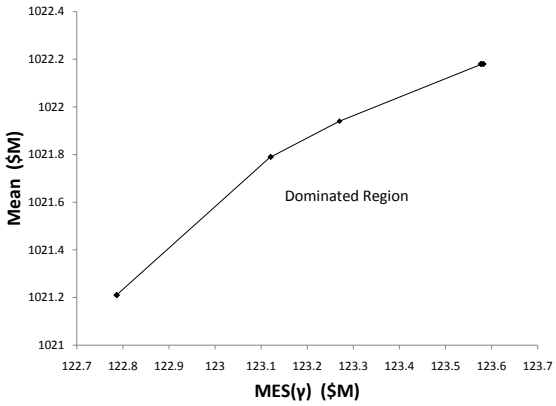


Figure 3.5: The absolute semideviation and the expected shortage mean-risk approximate efficient frontiers under the medium attack rate.

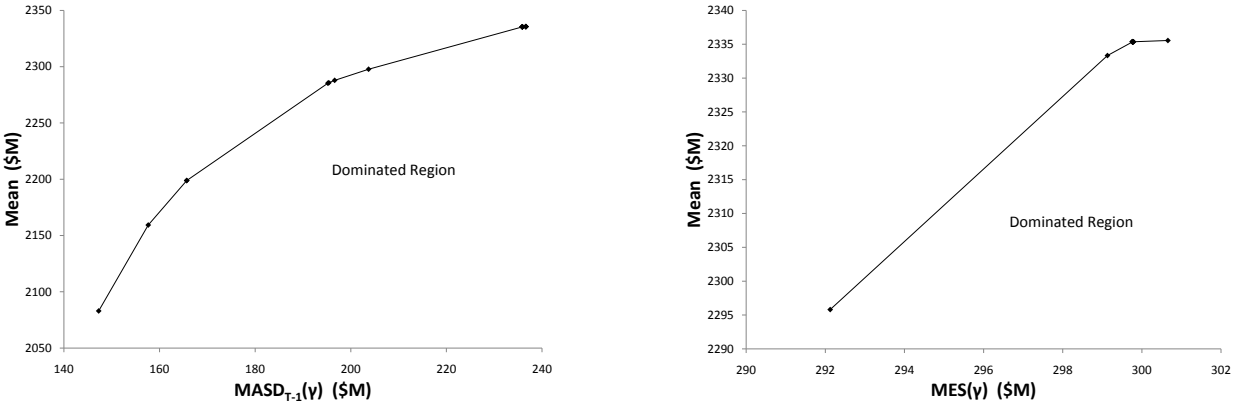


Figure 3.6: The absolute semideviation and the expected shortage mean-risk approximate efficient frontiers under the high attack rate.

4.0 OPTIMAL DESIGN OF THE ANNUAL INFLUENZA VACCINE WITH AUTONOMOUS MANUFACTURERS

In this chapter, we extend the multi-stage stochastic mixed-integer programming model of Chapter 3 to consider the hierarchical relationship between the Committee and the influenza vaccine manufacturers. This hierarchy results from the fact that the Committee optimizes the societal vaccination benefit by taking into account production decisions of the manufacturers, who maximize their own profits. The manufacturers' profit maximization problem is affected by the strain selection decisions of the Committee because not all strains grow well in chicken eggs. In return, the production decisions of the manufacturers affect the societal vaccination benefit because they determine the coverage rates and timely availability of the flu shot.

The government is not the only purchaser of the influenza vaccine. For example, 51% of the total flu shot produced was sold to physician offices and hospitals in the 2007-08 flu season [90]. The influenza vaccine price, unlike that of other vaccines, is not controlled by the government [51], and fluctuates in a competitive vaccine market. For example, in the 2000-01 flu season, manufacturing difficulties, which resulted in an overall delay of about 6-8 weeks in shipping the flu shot, created an initial shortage and a temporary price spike [73]. Currently, six manufacturers provide the influenza vaccine for the United States market [40]. Each manufacturer sets a production level to maximize its own profits by taking into account the decisions of the other manufacturers. Then, the production yield is realized, the actual quantity produced is brought to the market, and price emerges according to the total number of doses supplied by all manufacturers. We analyze the effects of yield uncertainty, price, and production cost on manufacturers' production decisions. We also analyze the properties of the consumer surplus using a hyperbolic demand function.

4.1 BILEVEL PROGRAMMING MODEL FORMULATION

We formulate a bilevel programming model for the annual influenza strain selection problem that incorporates both the Committee's and the vaccine manufacturers' perspectives. In the upper level, the Committee selects the strains to maximize the expected societal vaccination benefit considering the uncertainties in the prevalence of each strain and in the flu shot supply. In the lower level, each manufacturer sets production levels of the selected strains to maximize its profit based on the decisions of the other manufacturers as well as the uncertain production yield ratios and production costs. This model extends that of Chapter 3, as it could adjust production quantities of the manufacturers dynamically in response to selected strains.

Let $\mathcal{K}$ be the set of influenza vaccine manufacturers, and set $K = |\mathcal{K}|$. In addition, let $\mathcal{I}$ be the set of candidate strains for the influenza vaccine in a T-stage problem. Some portion of the uncertainty about the characteristics of the flu season unfolds at each time stage. The Committee selects the strains in the first $T-1$ stages. Manufacturers set their strain production levels after all strains are selected in stage $T-1$. Finally, all inherent uncertainty about the flu season is realized in the last stage, T.

As in Chapter 3, we model uncertainty using a scenario tree (see Figure 3.1) under a finite number of scenarios whose discrete probability distribution is independent of strain selection decisions [191]. However, in this chapter, scenarios model only the uncertainties in the prevalence of strains, while in Chapter 3 they also model the uncertainties in the production parameters, infection and shortage costs.

4.1.1 The Committee's Problem

We formulate the Committee's strain selection problem in the upper level as a multi–stage stochastic mixed–integer program. Denote by e_n^i the prevalence of strain $i \in \mathcal{I}$ at terminal node $n \in \mathcal{N}_T$, i.e. the ratio of the number of flu infections caused by strain i to the whole population. We represent the cost of a flu case by c. The shortage cost q captures the average cost of an unvaccinated person, e.g. health care costs, work loss cost, reduced productivity, to name a few. We assume that the society incurs this cost if the flu shot supply is less than a target level D.

We define binary variable $x_m^i = 1$ iff strain $i \in \mathcal{I}$ is selected at node $m \in \mathcal{N}_{t<T}$. For $\ell \in \mathcal{N}_{T-1}$, let $\mathcal{I}(\ell)$ be the set strains selected by the Committee, and set $\mathbf{X}_\ell \equiv (x_m^i)_{\mathcal{I}\times\mathscr{P}(\ell)}$. Denote by $v_{k\ell}^i$ the production level of strain i chosen by manufacturer k, and set $\mathbf{v}_{k\ell} \equiv (v_{k\ell}^i)_{\forall i\in\mathcal{I}(\ell)}$.

Let random variable $\beta^i \in (0,1)$ represent the production yield rate of strain $i \in \mathcal{I}$. Let $g_i(\cdot)$ and $G_i(\cdot)$ be the probability density function and the cumulative distribution function of β^i, respectively. We define $\bar{G}_i(\cdot) = 1 - G_i(\cdot)$. Moreover, denote by μ^i the mean of β^i. We assume that each manufacturer has the same yield distribution. Then, the expected flu shot production of manufacturer k at node $\ell \in \mathcal{N}_{T-1}$ is equal to $\mathbb{E}[\min_{i\in\mathcal{I}(\ell)}\{\beta^i v_{k\ell}^i\}]$. For $\ell \in \mathcal{N}_{T-1}$, denote by $W_\ell = \sum_{k\in\mathcal{K}} \mathbb{E}[\min_{i\in\mathcal{I}(\ell)}\{\beta^i v_{k\ell}^i\}]$, i.e. the expected flu shot production for terminal nodes $n \in \mathscr{T}(\ell) \cap \mathcal{N}_T$.

As described in Chapter 3, let $b(i,j)$ be the cross–effectiveness of strain i with strain j, and $\mathcal{H}(i,j) = \{h \in \mathcal{I} \mid b(h,j) > b(i,j)\}$ be the set of those strains that have higher cross–effectiveness with strain j than strain i does. We define nonnegative disjunctive variable $\delta_\ell^{ij} = W_\ell$ iff strain i but no strain from $\mathcal{H}(i,j)$ is selected.

Definition 4.1. *The "expected societal vaccination benefit" is the cost of infections averted by vaccinated people minus the cost of infections incurred by unvaccinated people in the target population. At node $\ell \in \mathcal{N}_{T-1}$, it is given by*

$$\Psi_\ell(\delta_\ell, W_\ell) = \sum_{n\in\mathscr{T}(\ell)\cap\mathcal{N}_T} \sum_{i,j\in\mathcal{I}} p_n c e_n^j \delta_\ell^{ij} b(i,j) - q\max\{D - W_\ell, 0\}. \tag{4.1}$$

We denote by $\Pi_{k\ell}$ the expected profit of manufacturer k at node $\ell \in \mathcal{N}_{T-1}$. Based on the notation defined above, the extensive form of the Committee's strain selection problem can be formulated as:

$$
\begin{aligned}
[\text{CP}] \qquad \max \quad & \sum_{\ell \in \mathcal{N}_{T-1}} p_\ell \Psi_\ell(\delta_\ell, W_\ell) && && (4.2)\\
\text{s.t.} \quad & \sum_{m \in \mathscr{P}(\ell)} x_m^i \leq 1 && \forall i \in \mathcal{I}, \ell \in \mathcal{N}_{T-1}, && (4.3)\\
& \delta_\ell^{ij} \leq M \min \left\{ 1 - \sum_{m \in \mathscr{P}(\ell)} x_m^h, \sum_{m \in \mathscr{P}(\ell)} x_m^i \right\} && \forall i,j \in \mathcal{I}, h \in \mathcal{H}(i,j), \ell \in \mathcal{N}_{T-1}, && (4.4)\\
& \mathbf{v}_{k\ell} \in \operatorname{argmax} \Pi_{k\ell} && \forall k \in \mathcal{K}, \ell \in \mathcal{N}_{T-1}, && (4.5)\\
& W_\ell = \sum_{k \in \mathcal{K}} \mathbb{E}[\min_{i \in \mathcal{I}(\ell)} \{\beta^i v_{k\ell}^i\}] && \ell \in \mathcal{N}_{T-1}, && (4.6)\\
& 0 \leq \delta_\ell^{ij} \leq W_\ell, x_\ell^i \in \{0,1\} && \forall i \in \mathcal{I}, \ell \in \mathcal{N}_{T-1}, m \in \mathcal{N}_{t<T}. && (4.7)
\end{aligned}
$$

Objective function (4.2) maximizes the expected societal vaccination benefit. We assume a risk-neutral decision maker; hence, the objective function is based on the expectation of all possible outcomes. Constraints (4.3) enforce each strain to be selected at most once under each scenario. In (4.4), M is a large positive number, e.g. the industry's overall flu shot production capacity. These constraints consider the antigenically closest flu shot strain to find the the immunity protection against a particular circulating flu virus [146]. Clearly, this modeling assumption underestimates the immunity protection provided by the flu shot. Finally, constraints (4.5) ensure that each manufacturer maximizes its expected profit. Note that $\mathbf{v}_{k\ell}$ depends on the equilibrium outcome of the Manufacturers' Problem.

4.1.2 Derivation of the Price Function

In this section, we derive the price function that is used in the profit function of the manufacturers. Consider $\mathcal{M}$ individuals, each of whom has a value ϑ for the influenza vaccine. This value depends on the price in addition to the likelihood of getting infected and the resulting costs of heath care, lost income, to name a few. Brito et al. [21], Deo and Corbett [56], and Cho [45] assumed that ϑ follows a uniform distribution on a constant range, which lead to

a linear demand function. We consider a variable consumer value range as a function of the influenza vaccine price ρ.

The value of the influenza vaccine for the high-risk population, e.g. the elderly, pregnant women, young children does not depend much on the price. However, people in the low-risk population, e.g. healthy adults, would take price into account more seriously. Let $\bar{\vartheta} > 0$ be the upper limit of ϑ. We assume that $\bar{\vartheta}$ is constant and does not vary with price. Denote by $\underline{\vartheta}(\rho)$ the lower limit of ϑ at price ρ. We assume that $\underline{\vartheta}(\rho)$ is decreasing in ρ.

Given the influenza vaccine price ρ, we assume that the consumer value ϑ follows a uniform distribution $\mathcal{F}_\rho(\cdot)$ in $[\underline{\vartheta}(\rho), \bar{\vartheta}]$. Note that the range $(\bar{\vartheta} - \underline{\vartheta}(\rho))$ is a nondecreasing function of ρ, as $\underline{\vartheta}(\rho)$ is decreasing in ρ. We assume that this function is defined by $\hbar\rho = (\bar{\vartheta} - \underline{\vartheta}(\rho))$ for $\hbar > 0$.

A rational consumer, with valuation ρ, is indifferent between getting vaccinated and not getting vaccinated. Hence, $Q = \mathcal{M}(1 - \mathcal{F}_\rho(\rho))$ is the total demand at that price. Substituting $1 - \mathcal{F}(\rho) = (\bar{\vartheta} - \rho)/\hbar\rho$, we get $Q = \mathcal{M}(\bar{\vartheta} - \rho)/\hbar\rho$. Setting $a = \bar{\vartheta}$ and $h = \hbar/\mathcal{M}$, the following hyperbolic price function is obtained:

$$\rho = \frac{a}{hQ + 1}. \tag{4.8}$$

Hyperbolic demand functions are well studied in economics [20, 110, 116], and have applications in electricity pricing [142, 192].

Deo and Corbett [56], and Cho [45] used a linear demand function $\rho = a - hQ$. Unlike those studies, we consider multiple flu shot strains. The revenue function $(a - hQ)Q$ is not concave for $Q = \min_i\{v^i\beta^i\}$ for a linear demand function. However, the hyperbolic function (4.8) preserves concavity as proven in Proposition 4.1(i). As a result, although a hyperbolic price function is somewhat more sophisticated than a linear one, it provides analytical tractability in our model.

4.1.3 The Manufacturers' Problem

We assume that the industry consists of K identical firms, which ensures that the equilibrium production quantity of each manufacturer must be the same as they also observe the same

yield distribution. For simplicity of the notation we drop the manufacturer index k, and the scenario tree node index ℓ hereafter in this section. In addition, we represent $\mathcal{I}(\ell)$ by I.

Let φ^i be the unit production cost of strain i. For constant parameters $a > 0$, $h > 0$, and the total flu shot supply Q, the flu shot price is equal to $\frac{a}{hQ+1}$, which is strictly decreasing in Q. Note that the expected unit production cost of the flu shot is equal to

$$\mathbb{E}[\sum_I \varphi^i/\beta^i] \geq \sum_I \varphi^i/\mu^i, \tag{4.9}$$

where the inequality follows from Jensen's inequality. If the reservation price a is less than or equal to $\sum_I \varphi^i/\mu^i$, then manufacturers' will not produce anything. Therefore, we assume that $a > \sum_I \varphi^i/\mu^i$. Moreover, for notational convenience we assume that $a/hK \geq 1$.

The manufacturer's problem can be formulated as:

$$\text{[MP]} \qquad \max\ \Pi = \mathbb{E}\left[\frac{a \min_I\{\beta^i v^i\}}{hK \min_I\{\beta^i v^i\} + 1}\right] - \sum_I \varphi^i v^i. \tag{4.10}$$

Proposition 4.1. *(i)* Π *is jointly concave in* $\mathbf{v}$.

(ii) The optimal solution $\mathbf{v}^*$ *is given by the unique solution to the first-order condition*

$$\frac{\partial \Pi}{\partial v^i} = \int_0^1 \frac{a\beta^i}{(hK\beta^i v^i + 1)^2} \prod_{i' \neq i} \bar{G}_{i'}\left(\beta^i v^i / v^{i'}\right) g_i(\beta^i) d\beta^i = \varphi^i \qquad i \in I. \tag{4.11}$$

Proof. *(i)* Π is the composition of strictly increasing and concave revenue function $aQ/(hQ+1)$ with $\min_I\{\beta^i v^i\}$. The result follows from the fact that $\min_I\{\beta^i v^i\}$ is also concave.

(ii) Rewrite the expected revenue as

$$\mathbb{E}\left[\frac{a \min_I\{\beta^i v^i\}}{hK \min_I\{\beta^i v^i\} + 1}\right] = \int_0^{a/hK} Pr\left(\frac{a \min_I\{\beta^i v^i\}}{hK \min_I\{\beta^i v^i\} + 1} \geq r\right) dr = \int_0^{a/hK} \prod_I \bar{G}_i\left(\frac{r/v^i}{a - hKr}\right) dr.$$

Using the change of variables $\hat{r} = \frac{r}{a - hKr}$,

$$\mathbb{E}\left[\frac{a \min_I\{\beta^i v^i\}}{hK \min_I\{\beta^i v^i\} + 1}\right] = \int_0^{a/hK} \frac{a}{(hK\hat{r} + 1)^2} \prod_I \bar{G}_i\left(\hat{r}/v^i\right) d\hat{r}.$$

The first-order optimality condition is given by

$$\frac{\partial \Pi}{\partial v^i} = \left(\frac{1}{v^i}\right)^2 \int_0^{a/hK} \frac{a\hat{r}}{(hK\hat{r} + 1)^2} \prod_{I \setminus \{i\}} \bar{G}_j\left(\hat{r}/v^j\right) g_i(\hat{r}/v^i) d\hat{r} - \varphi^i = 0 \qquad i \in I.$$

Using the change of variables $\beta^i = \hat{r}/v^i$,

$$\int_0^{a/hK} \frac{a\beta^i}{(hK\beta^i v^i + 1)^2} \prod_{I\setminus\{i\}} \bar{G}_j\left(\beta^i v^i / v^j\right) g_i(\beta^i) d\beta^i = \varphi^i \qquad i \in I.$$

Note that we can replace a/hK in the upper limit of the integral with one as the strain production yield is less than one. The left-hand side of (4.11) is the expected marginal revenue of increasing v^i, which is equal to $a\mu^i > \varphi^i$ for $v^i = 0$. This value is continuous, nonnegative and strictly decreasing in v^i, hence the optimal v^{i*} $i \in I$ is unique. □

Note that in the deterministic case when there is no yield uncertainty, the optimal production level of each strain must satisfy $v^i/\mu^i = w$ for $i \in I$, where w is a nonnegative scalar variable that represents the target flu shot production. The manufacturer's problem can then be reformulated as

$$\max \ \widehat{\Pi} = \frac{aw}{hKw + 1} - \sum_I \varphi^i \frac{w}{\mu^i}, \tag{4.12}$$

and the first-order optimality condition is given by

$$\frac{a}{(hKw + 1)^2} = \sum_I \frac{\varphi^i}{\mu^i}. \tag{4.13}$$

Condition (4.13) is both necessary and sufficient as $\widehat{\Pi}$ is concave in w. Let $\hat{w}$ be the optimal flu shot production level that maximizes $\widehat{\Pi}$. Denote by $\hat{v}^i = \hat{w}/\mu^i$ the optimal production level of strain $i \in I$. Deo and Corbett [56], and Cho [45] studied the effect of yield uncertainty on the flu shot production. They considered a single strain, and found that the yield uncertainty can reduce the industry output. Proposition 4.2 also considers a single flu shot strain, and compares the optimal production levels between the deterministic and stochastic cases.

Proposition 4.2. *If there is a single flu shot strain i, then $\hat{v}^i$, the optimal production level in the deterministic case, is no smaller than the optimal production level under uncertainty v^{i*}.*

Proof. If there is a single strain i, then condition (4.11) reduces to

$$\int_0^1 \frac{a\beta^i}{(hK\beta^i v^{i*}+1)^2} g_i(\beta^i) d\beta^i = \mathbb{E}\left[\frac{a\beta^i}{(hK\beta^i v^{i*}+1)^2}\right] = \varphi^i, \tag{4.14}$$

and condition (4.13) reduces to

$$\frac{a\mu^i}{(hK\mu^i \hat{v}^i+1)^2} = \varphi^i. \tag{4.15}$$

Using (4.14), (4.15), and Jensen's inequality

$$\frac{a\mu^i}{(hK\mu^i \hat{v}^i+1)^2} = \mathbb{E}\left[\frac{a\beta^i}{(hK\beta^i v^{i*}+1)^2}\right] \leq \frac{a\mu^i}{(hK\mu^i v^{i*}+1)^2}. \tag{4.16}$$

□

Next, we investigate the effect of production yield uncertainty on the optimal strain production levels when there are multiple strains in the flu shot. To gain additional insights, we analyze a special case in which only a particular strain $j \in I$ has yield uncertainty. In this case, the optimal production level of each flu shot strain $i \neq j$ with certain yield must be equal as the amount of the flu shot produced is limited by the lowest strain production quantity. To ensure this condition we replace $\mu^i v^i$ $i \in I \setminus \{j\}$ by a nonnegative scalar variable z that represents the target flu shot production. The manufacturer's problem can then be reformulated as

$$\max \ \widetilde{\Pi} = \mathbb{E}\left[\frac{a \min\{z, \beta^j v^j\}}{hK \min\{z, \beta^j v^j\}+1}\right] - z \sum_{I\setminus\{j\}} \frac{\varphi^i}{\mu^i} - \varphi^j v^j. \tag{4.17}$$

We denote optimal solution that maximizes $\widetilde{\Pi}$ by $(\tilde{v}^j, \tilde{z})$. Recall that Proposition 4.2 shows that $v^{i*} \leq \hat{v}^i$, when i is the only flu shot strain. Theorem 4.1 states that when there are multiple strains in the flu shot, the optimal strain production levels may be either less or more than those in the deterministic case depending on the reservation price a, even if only one of the strains has uncertain production yield. Before presenting this result, we prove some technical results.

Proposition 4.3. *$\tilde{v}^j$ and $\tilde{z}$ are strictly increasing in a.*

Proof. Without loss of generality let z be the production level of a strain whose yield is equal to one. Then, Proposition 4.1 is valid for $\widetilde{\Pi}$. From condition (4.11),

$$Q_1 = \frac{\partial \widetilde{\Pi}}{\partial v^j} = \int_0^{z/v^j} \frac{a\beta^j}{(hK\beta^j v^j + 1)^2} g_j(\beta^j) d\beta^j - \varphi^j = 0, \tag{4.18a}$$

$$Q_2 = \frac{\partial \widetilde{\Pi}}{\partial z} = \frac{a}{(hKz+1)^2} \bar{G}_j(z/v^j) - \sum_{I\setminus\{j\}} \frac{\varphi^i}{\mu^i} = 0. \tag{4.18b}$$

Next, we apply the Implicit Function Theorem [161] to obtain

$$\frac{\partial Q_1}{\partial \tilde{v}^j}\frac{\partial \tilde{v}^j}{\partial a} + \frac{\partial Q_1}{\partial \tilde{z}}\frac{\partial \tilde{z}}{\partial a} = -\frac{\partial Q_1}{\partial a}, \tag{4.19a}$$

$$\frac{\partial Q_2}{\partial \tilde{v}^j}\frac{\partial \tilde{v}^j}{\partial a} + \frac{\partial Q_2}{\partial \tilde{z}}\frac{\partial \tilde{z}}{\partial a} = -\frac{\partial Q_2}{\partial a}. \tag{4.19b}$$

Note that $\frac{\partial Q_1}{\partial \tilde{v}^j} < 0$, $\frac{\partial Q_1}{\partial \tilde{z}} > 0$, $\frac{\partial Q_1}{\partial a} > 0$, $\frac{\partial Q_2}{\partial \tilde{v}^j} > 0$, $\frac{\partial Q_2}{\partial \tilde{z}} < 0$, and $\frac{\partial Q_2}{\partial a} > 0$. Therefore, $\frac{\partial \tilde{v}^j}{\partial a}$ and $\frac{\partial \tilde{z}}{\partial a}$ must have the same sign. We solve (4.19a) and (4.19b) simultaneously to obtain

$$\left(\frac{\partial Q_1}{\partial \tilde{v}^j}\frac{\partial Q_2}{\partial \tilde{z}} - \frac{\partial Q_1}{\partial \tilde{z}}\frac{\partial Q_2}{\partial \tilde{v}^j}\right)\frac{\partial \tilde{v}^j}{\partial a} = \frac{\partial Q_1}{\partial \tilde{z}}\frac{\partial Q_2}{\partial a} - \frac{\partial Q_1}{\partial a}\frac{\partial Q_2}{\partial \tilde{z}}. \tag{4.20}$$

The right-hand side of (4.20) is positive. The expression in the brackets in the left-hand side is equal to $\det(\partial^2 \Pi)$, which is nonnegative due to Proposition 4.1 (i). As a result, $\frac{\partial \tilde{v}^j}{\partial a} > 0$, so that $\frac{\partial \tilde{z}}{\partial a} > 0$, as it must have the same sign. □

Proposition 4.3 states that a higher reservation price a will provide the manufacturers' incentive to produce more flu shot strains with deterministic or uncertain yield.

Lemma 4.1. *For a given reservation price a, let*

$$L(a) = \frac{\partial \widetilde{\Pi}(\hat{w}, \gamma\hat{w})}{\partial z} = \frac{a}{(hK\gamma\hat{w}+1)^2} \bar{G}_j(\gamma) - \sum_{I\setminus\{j\}} \frac{\varphi^i}{\mu^i}, \tag{4.21}$$

where γ satisfies

$$\int_0^{\gamma} \frac{a\beta^j}{(hK\beta^j\hat{w}+1)^2} g_j(\beta^j) d\beta^j = \varphi^j. \tag{4.22}$$

Then the following results hold

(i) $0 < \gamma < 1$,

(ii) γ *is strictly decreasing in* a,

(iii) $L(a)$ *is strictly increasing in* a.

Proof. (i) For $\gamma = 0$ the left-hand side of (4.22) is zero, and for $\gamma = 1$,

$$\int_0^1 \frac{a\beta^j}{(hK\beta^j\hat{w}+1)^2} g_j(\beta^j)d\beta^j \geq \int_0^1 \frac{a\beta^j}{(hK\hat{w}+1)^2} g_j(\beta^j)d\beta^j = \frac{a\mu^j}{(hK\hat{w}+1)^2} = \varphi^j + \mu^j \sum_{I\setminus\{j\}} \frac{\varphi^i}{\mu^i},$$

where the second equality follows from (4.13). Observe that (4.22) is strictly increasing and continuous in γ, so $0 < \gamma < 1$.

(ii) Note that γ is a function of a, but we suppress this for notational simplicity. Differentiate both sides of (4.13) with respect to a,

$$\frac{(hK\hat{w}+1)^2 - 2hKa(hK\hat{w}+1)\frac{\partial\hat{w}}{\partial a}}{(hK\hat{w}+1)^4} = 0 \Rightarrow \frac{\partial\hat{w}}{\partial a} = \frac{hK\hat{w}+1}{2ahK}. \tag{4.23}$$

Now, differentiate both sides of (4.22) with respect to a,

$$\frac{\partial\gamma}{\partial a}\frac{a\gamma}{(hK\gamma\hat{w}+1)^2} g_j(\gamma) + \int_0^\gamma \frac{(hK\beta^j\hat{w}+1) - 2ahK\beta^j\frac{\partial\hat{w}}{\partial a}}{(hK\beta^j\hat{w}+1)^3}\beta^j g_j(\beta^j)d\beta^j = 0. \tag{4.24}$$

Using (4.23), we can expand the second term in (4.24) as,

$$\int_0^\gamma \frac{(hK\beta^j\hat{w}+1) - (hK\beta^j\hat{w}+\beta^j)}{(hK\beta^j\hat{w}+1)^3}\beta^j g_j(\beta^j)d\beta^j > 0.$$

Thus, it follows that $\frac{\partial\gamma}{\partial a} < 0$.

(iii) Differentiate $L(a)$ with respect to a to obtain

$$\begin{aligned}\frac{\partial L(a)}{\partial a} =& \frac{(hK\gamma\hat{w}+1) - 2ahK(\gamma\frac{\partial\hat{w}}{\partial a} + \frac{\partial\gamma}{\partial a}\hat{w})}{(hK\gamma\hat{w}+1)^3}\bar{G}_j(\gamma) \\ =& \frac{(hK\gamma\hat{w}+1) - (hK\gamma\hat{w}+\gamma) - 2ahK\frac{\partial\gamma}{\partial a}\hat{w}}{(hK\gamma\hat{w}+1)^3}\bar{G}_j(\gamma) > 0,\end{aligned} \tag{4.25}$$

where the first inequality follows from (4.23). The last inequality follows since $\gamma < 1$ and $\frac{\partial\gamma}{\partial a} < 0$. □

Lemma 4.2. *For $0 < u \leq 1$ and reservation price a, let*

$$B(a,u) = \frac{\partial \widetilde{\Pi}(\grave{v}^j, u\grave{v}^j)}{\partial z} = \frac{a}{(hKu\grave{v}^j + 1)^2}\bar{G}_j(u) - \sum_{I\setminus\{j\}} \frac{\varphi^i}{\mu^i}, \tag{4.26}$$

where $\grave{v}^j \geq 0$ satisfies,

$$\frac{\partial \widetilde{\Pi}(\grave{v}^j, u\grave{v}^j)}{\partial v^j} = \int_0^u \frac{a\beta^j}{(hK\beta^j\grave{v}^j + 1)^2} g_j(\beta^j)d\beta^j = \varphi^j. \tag{4.27}$$

Then $B(a,u)$ is strictly decreasing in a and u.

Proof. Note that $\grave{v}$ is a function of a, but we suppress it for notational simplicity. Differentiate both sides of (4.27) with respect to a to obtain

$$\begin{aligned} &\int_0^u \frac{(hK\beta^j\grave{v}^j + 1)^2 - 2hKa\beta^j(hK\beta^j\grave{v}^j + 1)\frac{\partial \grave{v}^j}{\partial a}}{(hK\beta^j\grave{v}^j + 1)^4}\beta^j g_j(\beta^j)d\beta^j = 0 \\ \Rightarrow &\frac{\partial \grave{v}^j}{\partial a} = \frac{\int_0^u \frac{\beta^j}{(hK\beta^j\grave{v}^j+1)^2} g_j(\beta^j)d\beta^j}{\int_0^u \frac{2hKa(\beta^j)^2}{(hK\beta^j\grave{v}^j+1)^3} g_j(\beta^j)d\beta^j} = \frac{\varphi^j/a}{2hKa\int_0^u \frac{(\beta^j)^2}{(hK\beta^j\grave{v}^j+1)^3} g_j(\beta^j)d\beta^j}, \end{aligned} \tag{4.28}$$

where the equality follows from (4.27). Observe that in the denominator of (4.28)

$$\int_0^u \frac{(\beta^j)^2}{(hK\beta^j\grave{v}^j + 1)^3} g_j(\beta^j)d\beta^j < \frac{u}{(hKu\grave{v}^j + 1)} \int_0^u \frac{\beta^j}{(hK\beta^j\grave{v}^j + 1)^2} g_j(\beta^j)d\beta^j = \frac{u}{(hKu\grave{v}^j + 1)}\frac{\varphi^j}{a}. \tag{4.29}$$

Using inequality (4.29) in (4.28), we obtain

$$\frac{\partial \grave{v}^j}{\partial a} > \frac{\varphi^j/a}{2hKa\frac{u}{(hKu\grave{v}^j+1)}\frac{\varphi^j}{a}} = \frac{(hKu\grave{v}^j + 1)}{2hKau}. \tag{4.30}$$

Differentiate $B(a,u)$ with respect to a to obtain

$$\frac{\partial B(a,u)}{\partial a} = \frac{(hKu\grave{v}^j + 1) - 2hKau\frac{\partial \grave{v}^j}{\partial a}}{(hKu\grave{v}^j + 1)^3}\bar{G}_j(u) < 0, \tag{4.31}$$

where the inequality follows from (4.30). Thus, $B(a,u)$ is strictly decreasing in a. Observe from (4.27) that $\grave{v}^j$ is strictly increasing in u. It follows from (4.26) that $B(a,u)$ is strictly decreasing in u.

□

Lemma 4.3. *Let $\grave{u}$ be such that $\int_0^{\grave{u}} a\beta^j g_j(\beta_j)d\beta^j = \varphi^j$, and $B(a,u)$ be as in Lemma 4.2.*

(i) $B(a,\grave{u}) \geq 0$

(ii) There exists a unique u, $\grave{u} \leq u \leq 1$, denoted by $u_1(a)$, such that $B(a,u_1(a)) = 0$. Moreover, $u_1(a)$ is strictly decreasing in a.

Proof. (i) From (4.18b)

$$\int_0^{\tilde{u}} \frac{a\beta^j}{(hK\beta^j\tilde{v}^j+1)^2} g_j(\beta^j)d\beta^j = \varphi^j = \int_0^{\grave{u}} a\beta^j g_j(\beta_j)d\beta^j. \tag{4.32}$$

Thus, $\grave{u} \leq \tilde{u}$. Furthermore, it follows from (4.27) that $\grave{v}^j = 0$ for $u = \grave{u}$. Thus,

$$B(a,\grave{u}) = a\bar{G}_j(\grave{u}) - \sum_{I\setminus\{j\}} \frac{\varphi^i}{\mu^i} \geq a\bar{G}_j(\tilde{u}) - \sum_{I\setminus\{j\}} \frac{\varphi^i}{\mu^i} \geq \frac{a}{(hK\tilde{z}+1)^2}\bar{G}_j(\tilde{u}) - \sum_{I\setminus\{j\}} \frac{\varphi^i}{\mu^i} = 0. \tag{4.33}$$

(ii) Clearly, $B(a,1) = -\sum_{I\setminus\{j\}} \varphi^i/\mu^i < 0$. Lemma 4.2 (ii) states that $B(a,u)$ is strictly increasing in u, and we showed that $B(a,\grave{u}) \geq 0$ in the first part. Thus, there must exists some unique u between $\grave{u}$ and 1, denoted by $u_1(a)$, such that $B(a,u_1(a)) = 0$. As $B(a,u)$ is strictly decreasing in a and u from Lemma 4.2 (ii), $u_1(a)$ is strictly decreasing in a. □

Proposition 4.4. *There exists an $a_1 > \sum_I \varphi^i/\mu^i$ such that*

$$\begin{cases} \tilde{v}^j = \hat{w} & \text{if } a = a_1, \\ \tilde{v}^j < \hat{w} & \text{if } a < a_1, \\ \tilde{v}^j > \hat{w} & \text{if } a > a_1. \end{cases}$$

Proof. We first show that there exists an a_1 such that $\tilde{v}^j = \hat{w}$ when $a = a_1$. From (4.22), it is clear that $(v^j, z) = (\hat{w}, \gamma\hat{w})$ satisfies (4.18a) for any $a > \sum_I \varphi^i/\mu^i$. It simplifies to show that $(\hat{w}, \gamma\hat{w})$ also satisfies (4.18b) for some $a_1 > \sum_I \varphi^i/\mu^i$. Note that for $a = \sum_I \varphi^i/\mu^i$ we have $\hat{w} = 0$ from (4.13). Moreover, it follows from (4.22) that

$$\int_0^{\gamma} \beta^j g_j(\beta^j)d\beta^j = \mathbb{E}[\beta^j|\beta^j \leq \gamma]G_j(\gamma) = \frac{\varphi^j}{\sum_I \varphi^i/\mu^i} \Rightarrow G_j(\gamma) \geq \frac{\varphi^j/\mu^j}{\sum_I \varphi^i/\mu^i}. \tag{4.34}$$

Then,

$$L\left(\sum_I \varphi^i/\mu^i\right) = \frac{\varphi^j}{\mu^j} - G_j(\gamma)\sum_I \frac{\varphi^i}{\mu^i} \le 0. \tag{4.35}$$

Since $L(a)$ is strictly increasing in a, there must exist some $a_1 > \sum_I \varphi^i/\mu^i$ such that $L(a_1) = 0$, or equivalently, $(\hat{w}, \gamma\hat{w})$ satisfies (4.18b) when $a = a_1$.

Next, we show that $\tilde{v}^j > \hat{w}$ when $a > a_1$. Consider any fixed $\check{a} > a_1$. Let $(\check{v}^j, \check{z})$ be the corresponding optimal solution. Then, from Lemma 4.3 (ii), $u_1(\check{a}) < u_1(a_1)$, which implies that $\bar{G}_j(u_1(\check{a})) > \bar{G}_j(u_1(a_1))$. From (4.18b) we have that

$$\frac{\check{a}}{(hKu_1(\check{a})\check{v}^j + 1)^2}\bar{G}_j(u_1(\check{a})) = \frac{a_1}{(hKu_1(a_1)\hat{w} + 1)^2}\bar{G}_j(u_1(a_1)).$$

Thus, $\check{v}^j > \hat{w}$. Similarly, we can show that $\check{v}^j < \hat{w}$ when $\check{a} < a_1$. □

Lemma 4.4. *For $0 < u \le 1$ and reservation price a, parameterize the upper integrand of (4.18a) by u. Let*

$$J(a,u) = \frac{\partial\widetilde{\Pi}(\bar{v}^j, \bar{z})}{\partial v^j} = \int_0^u \frac{a\beta^j}{(hK\beta^j\frac{\bar{z}}{u} + 1)^2} g_j(\beta^j)d\beta^j - \varphi^j, \tag{4.36}$$

where $\bar{z}$ satisfies

$$(hK\bar{z} + 1)^2 = \frac{a}{\sum_{I\setminus\{j\}} \varphi^i/\mu^i}\bar{G}_j(u). \tag{4.37}$$

Then $J(a,u)$ is strictly increasing in a and u.

Proof. Note that $\bar{z}$ is a function of a. Differentiate both sides of (4.37) with respect to a,

$$\frac{\partial\bar{z}}{\partial a} = \frac{\bar{G}_j(u)}{\sum_{I\setminus\{j\}} \varphi^i/\mu^i}\frac{(hK\bar{z} + 1)}{2hK(hK\bar{z} + 1)^2} = \frac{(hK\bar{z} + 1)}{2hKa}, \tag{4.38}$$

where the second equality follows from (4.37).

Differentiating $J(a,u)$ with respect to a,

$$\begin{aligned}\frac{\partial J(a,u)}{\partial a} &= \int_0^u \left(\frac{(hK\beta^j\frac{\bar{z}}{u} + 1)^2 - 2hKa\frac{\beta^j}{u}(hK\beta^j\frac{\bar{z}}{u} + 1)\frac{\partial\bar{z}}{\partial a}}{(hK\beta^j\frac{\bar{w}}{u} + 1)^4}\right)\beta^j g_j(\beta^j)d\beta^j \\ &= \int_0^u \left(\frac{(hK\beta^j\frac{\bar{z}}{u} + 1)^2 - (hK\beta^j\frac{\bar{z}}{u} + 1)(hK\beta^j\frac{\bar{z}}{u} + \frac{\beta^j}{u})}{(hK\beta^j\frac{\bar{z}}{u} + 1)^4}\right)\beta^j g_j(\beta^j)d\beta^j > 0,\end{aligned} \tag{4.39}$$

where the first equality follows from (4.38), and the inequality follows since $\beta^j \le u$. Therefore, $J(a,u)$ is strictly increasing in a. Observe from (4.37) that $\bar{z}$ is strictly decreasing in u. It follows from (4.36) that $J(a,u)$ is strictly increasing in u. □

Lemma 4.5. *Let $\bar{u}$ be such that $\bar{G}_j(\bar{u}) = \frac{\sum_{I\setminus\{j\}} \varphi^i/\mu^i}{a}$.*

(i) $J(a, \bar{u}) \geq 0$.

(ii) There exists a unique u, $0 \leq u \leq \bar{u}$, denoted by $u_2(a)$, such that $J(a, u_2(a)) = 0$. Moreover, $u_2(a)$ is strictly decreasing in a.

Proof. *(i)* Let $\tilde{u} = \tilde{z}/\tilde{v}^j$. It follows from (4.18b) that

$$\bar{G}_j(\tilde{u}) = \frac{(hK\tilde{z}+1)^2 \sum_{I\setminus\{j\}} \varphi^i/\mu^i}{a} \geq \frac{\sum_{I\setminus\{j\}} \varphi^i/\mu^i}{a} = \bar{G}_j(\bar{u}). \tag{4.40}$$

Thus, $\bar{u} \geq \tilde{u}$. Furthermore, it follows from (4.37) that $\bar{z} = 0$ for $u = \bar{u}$. Thus,

$$J(a,\bar{u}) = \int_0^{\bar{u}} a\beta^j g_j(\beta^j) d\beta^j - \varphi^j \geq \int_0^{\tilde{u}} a\beta^j g_j(\beta^j) d\beta^j - \varphi^j \geq \int_0^{\tilde{u}} \frac{a\beta^j}{(hK\beta^j\tilde{v}^j+1)^2} g_j(\beta^j) d\beta^j - \varphi^j = 0,$$

where the first inequality follows from (4.40), and the last equality follows from (4.18a).

(ii) Clearly, $\lim_{u\to 0^+} J(a,u) = -\varphi^j < 0$. Lemma 4.4 states that $J(a,u)$ is strictly increasing in u, and we showed that $J(a,\bar{u}) \geq 0$ in the first part. Thus, there must exists some unique u between 0 and $\bar{u}$, denoted by $u_2(a)$, such that $J(a, u_2(a)) = 0$. As $J(a,u)$ is strictly increasing in a and u from Lemma 4.4, $u_2(a)$ is strictly decreasing in a.

□

Proposition 4.5. *There exists an $a_2 > \sum_I \varphi^i/\mu^i$ such that*

$$\begin{cases} \tilde{z} = \hat{w} & \text{if } a = a_2, \\ \tilde{z} < \hat{w} & \text{if } a < a_2, \\ \tilde{z} > \hat{w} & \text{if } a > a_2. \end{cases}$$

Proof. We first show that there exists some a_2 such that $\tilde{z} = \hat{w}$ when $a = a_2$. Let $\eta = G_j^{-1}\left(\frac{\varphi^j/\mu^j}{\sum_I \varphi^i/\mu^i}\right)$. Recalling that $\hat{w}$ satisfies (4.13), it is straightforward to show that $(v^j, z) = (\frac{\hat{w}}{\eta}, \hat{w})$ satisfies (4.18b) for any $a > \sum_I \varphi^i/\mu^i$. We claim that there exists an a_2 such that $(\frac{\hat{w}}{\eta}, \hat{w})$ also satisfies (4.18a).

Note that for $a = \sum_I \varphi^i/\mu^i$ and $u = \eta$, it follows from (4.37) that

$$(hK\bar{z} + 1)^2 = \frac{\sum_I \varphi^i/\mu^i}{\sum_{I\setminus\{j\}} \varphi^i/\mu^i}\left(1 - \frac{\varphi^j/\mu^j}{\sum_I \varphi^i/\mu^i}\right) = 1. \tag{4.41}$$

Thus, $\bar{z} = 0$ for $a = \sum_I \varphi^i/\mu^i$ and $u = \eta$. Then from (4.36),

$$J(\sum_I \varphi^i/\mu^i, \eta) = \sum_I \frac{\varphi^i}{\mu^i}\int_0^\eta \beta^j g_j(\beta^j)d\beta^j - \varphi^j \le \sum_I \frac{\varphi^i}{\mu^i} G_j(\eta)\mu^j - \varphi^j = 0, \tag{4.42}$$

where the inequality follows from $\int_0^\eta \beta^j g_j(\beta^j)d\beta^j = G_j(\eta)\mathbb{E}[\beta^j|\beta^j \le \eta] \le G_j(\eta)\mu^j$, and the second equality follows from the definition of η. Since, $J(a, \eta)$ is strictly increasing in a from Lemma 4.4, there must exist some $a_2 > \sum_I \varphi^i/\mu^i$ such that $J(a_2, \eta) = 0$, or equivalently, $(v^j, z) = (\frac{\hat{w}}{\eta}, \hat{w})$ satisfies (4.18a) when $a = a_2$.

Next, we show that $\tilde{z} > \hat{w}$ for $a > a_2$. Consider any fixed $\breve{a} > a_2$. Let $(\breve{v}^j, \breve{z})$ be corresponding optimal solution of (4.17). Then, from Lemma 4.5 (ii), $u_2(\breve{a}) < u_2(a_2)$, which implies that $\bar{G}_j(u_2(\breve{a})) > \bar{G}_j(u_2(a_2))$. From (4.18b) we have that

$$(hK\breve{z} + 1)^2 = \frac{\breve{a}}{\sum_{I\setminus\{j\}} \varphi^i/\mu^i}\bar{G}_j(u_2(\breve{a})) > \frac{a_2}{\sum_{I\setminus\{j\}} \varphi^i/\mu^i}\bar{G}_j(u_2(a_2)) = (hK\hat{w} + 1)^2. \tag{4.43}$$

Thus, $\breve{z} > \hat{w}$. Similarly, we can show that $\breve{z} < \hat{w}$ when $\breve{a} < a_2$. □

Theorem 4.1. *There exists a_1 and a_2 such that $a_1 \le a_2$ and*

$$\begin{cases} \tilde{z} \le \tilde{v}^j < \hat{w} & \text{if } a < a_1, \\ \tilde{z} \le \hat{w} \le \tilde{v}^j & \text{if } a_1 \le a \le a_2, \\ \hat{w} < \tilde{z} \le \tilde{v}^j & \text{if } a > a_2. \end{cases}$$

Proof. The flu shot production is equal to $\min\{\tilde{z}, \beta^j \tilde{v}^j\}$. Observe that if $\tilde{z} > \tilde{v}^j$, we can decrease $\tilde{z}$ to $\tilde{v}^j$, and obtain the same flu shot production with less production cost. Hence, $\tilde{z} \leq \tilde{v}^j$ in any optimal solution of the manufacturers' problem. It follows from Proposition 4.4 that $\tilde{z} \leq \tilde{v}^j < \hat{w}$ when $a < a_1$, which shows the first part. Moreover, it follows from Proposition 4.5 that $\hat{w} < \tilde{z} \leq \tilde{v}^j$ when $a > a_2$, which shows the last part. Finally, when $a_1 \leq a \leq a_2$, we have $\tilde{v}^j \geq \hat{w}$ from Proposition 4.4, and $\tilde{z} \leq \hat{w}$ from Proposition 4.5, which shows the second part. □

Theorem 4.1 demonstrates the impact of yield uncertainty of one strain on the production levels of the other strains. There are two basic factors that influence the optimal production quantity of strain j facing yield uncertainty. First, yield uncertainty increases the effective unit production cost, i.e. $\frac{\varphi^j}{\beta^j} \geq \varphi^j$, so we need to reduce the production level of strain j due to the higher production cost. Second, we need to inflate the production level of strain j to account for the underlying yield uncertainty. These two factors act against each other. When the flu shot price is low, the first factor is dominant and the manufacturer needs to decrease the production level. When the flu shot price is high, the second factor becomes dominant and the manufacturer needs to increase the production level.

The more surprising result is that the optimal production level of the other strains could also increase or decrease as compared to the deterministic case. As discussed above, yield uncertainty effectively increases the unit cost of strain j and thus decreases the profit margin. Then, it seems intuitive that $\bar{z}$ should decrease as a result because there is no uncertainty in the yield of strains $i \neq j$, and the manufacturer need not inflate their production levels. Theorem 4.1 shows that this intuition is not necessarily true when the flu shot price is high. Indeed, when a is large enough ($a > a_2$), $\bar{z}$ is greater than $\hat{w}$, the production level when there is no uncertainty. The high flu shot price provides the manufacturer sufficient incentive to increase the production levels of strains $i \neq j$ with certain yield to match the corresponding larger production level of strain j in hope of a high yield realization.

Note that Theorem 4.1 characterizes the behavior of the strain production levels under uncertainty and in the deterministic case with respect to the reservation price. Next, we consider the expected flu shot production, again when only strain j has yield uncertainty.

Proposition 4.6 shows that the expected production under uncertainty can exceed the production in the deterministic case for large enough reservation price a.

Proposition 4.6. *There exists an $a_3 > a_2$ such that*

$$\begin{cases} \mathbb{E}[\min\{\tilde{z}, \beta^j \tilde{v}^j\}] = \hat{w} & \text{if } a = a_3, \\ \mathbb{E}[\min\{\tilde{z}, \beta^j \tilde{v}^j\}] < \hat{w} & \text{if } a < a_3, \\ \mathbb{E}[\min\{\tilde{z}, \beta^j \tilde{v}^j\}] > \hat{w} & \text{if } a > a_3. \end{cases}$$

Proof. We first show that $\mathbb{E}[\min\{\tilde{z}, \beta^j \tilde{v}^j\}] < \hat{w}$ if $a < a_2$. Observe that $\mathbb{E}[\min\{\tilde{z}, \beta^j \tilde{v}^j\}] \leq \tilde{z}$. It follows from Proposition 4.5 that $\tilde{z} < \hat{w}$ if $a < a_2$.

Next, we show that $\mathbb{E}[\min\{\tilde{z}, \beta^j \tilde{v}^j\}] > \hat{w}$ if $a > \frac{a_2}{(\mu^j)^2}$. Observe that $\mathbb{E}[\min\{\tilde{z}, \beta^j \tilde{v}^j\}] \geq \mathbb{E}[\min\{\tilde{z}, \beta^j \tilde{z}\}] = \mu^j \tilde{z}$. If $a > \frac{a_2}{(\mu^j)^2}$, it follows from (4.43) that

$$\begin{aligned} (hK\tilde{z}+1)^2 &= \frac{a}{\sum_{I\setminus\{j\}} \varphi^i/\mu^i} \bar{G}_j(u_2(a)) > \frac{a_2/(\mu^j)^2}{\sum_{I\setminus\{j\}} \varphi^i/\mu^i} \bar{G}_j(u_2\left(a_2/(\mu^j)^2\right)) \\ > \frac{a_2/(\mu^j)^2}{\sum_{I\setminus\{j\}} \varphi^i/\mu^i} \bar{G}_j(u_2(a_2)) &= \frac{(hK\hat{w}+1)^2}{(\mu^j)^2} > \frac{(hK\hat{w}+\mu^j)^2}{(\mu^j)^2} = (hK\frac{\hat{w}}{\mu^j}+1)^2. \end{aligned} \tag{4.44}$$

Thus, $\tilde{z} > \frac{\hat{w}}{\mu^j}$ if $a > \frac{a_2}{(\mu^j)^2}$, and so $\mathbb{E}[\min\{\tilde{z}, \beta^j \tilde{z}\}] = \mu^j \tilde{z} > \hat{w}$.

Finally, note that both $\mathbb{E}[\min\{\tilde{z}, \beta^j \tilde{v}^j\}]$ and $\hat{w}$ are nondecreasing continuous functions of a. Therefore, there exists a reservation price between a_2 and $\frac{a_2}{(\mu^j)^2}$ such that $\mathbb{E}[\min\{\tilde{z}, \beta^j \tilde{v}^j\}] = \hat{w}$. □

4.2 CONSUMER SURPLUS

Consumer surplus (CS) is an economic measure of consumer satisfaction, which is defined as the difference between what consumers are willing to pay for a good or service and its market

price. Given the total flu shot production Q, the consumer utility is equal to $\int_0^Q \frac{a}{hr+1}dr$, and the payment by the consumers is equal to $\frac{aQ}{hQ+1}$. Hence, the consumer surplus is given by

$$\text{CS} = \int_0^Q \frac{a}{hr+1}dr - \frac{aQ}{hQ+1} = \frac{a}{h}\ln(hQ+1) - \frac{aQ}{hQ+1}. \tag{4.45}$$

For fixed a, CS is increasing in Q as

$$\frac{\partial \text{CS}}{\partial Q} = \frac{a}{hQ+1} - \frac{a}{(hQ+1)^2} > 0. \tag{4.46}$$

Clearly, CS is increasing in a for fixed Q. Furthermore, Q is also increasing in a, and so CS is increasing in a. Deo and Corbett [56] found that the expected consumer surplus in an equilibrium is always smaller under yield uncertainty than that in the deterministic case, because fewer units are brought to the market when there is yield uncertainty. As a result of Proposition 4.6 and the fact that consumer surplus is increasing in the amount of the flu shot brought to the market, we conclude that the expected consumer surplus in an equilibrium is not always smaller under yield uncertainty than that in the deterministic case. We illustrate this phenomenon in our numerical experiments.

4.3 SOLUTION APPROACH

The Committee's Problem (CP) must consider the equilibrium of the Manufacturers' Problem (MP), hence it can not be solved by traditional multi-stage stochastic programming techniques. In Section 4.3.1, we use a multidimensional root finding technique to solve the MP. Then, in Section 4.3.2, we develop a branch-and-price algorithm to solve the CP based on its [50] reformulation.

4.3.1 Solving the Manufacturers' Problem

The first-order optimality conditions (4.11) form a nonlinear system of equations in $|\mathcal{I}(\ell)|$ dimensions. There are no bracketing methods available for multidimensional systems. All algorithms proceed from an initial guess using a variant of the Newton iteration.

We solve (4.11) using a modified version of Powell's Hybrid method [152]. The algorithm takes the Jacobian matrix as input, and utilizes a generalized trust region to keep each step under control. It retains the fast convergence of Newton's method, and reduces the residual when Newton's method is unreliable (see Section 4.5.1).

The diagonal and off-diagonal elements of the Jacobian matrix of the nonlinear system (4.11) are given by

$$\frac{\partial^2 \Pi}{\partial (v^i)^2} = \int_0^1 \frac{-2hKa(\beta^i)^2}{(hK\beta^i v^i + 1)^3} \prod_{i' \neq i} \bar{G}_{i'} \left(\beta^i v^i / v^{i'} \right) g_i(\beta^i) d\beta^i$$
$$- \int_0^1 \sum_{i' \neq i} \frac{a(\beta^i)^2}{(hK\beta^i v^i + 1)^2 v^{i'}} \prod_{i'' \neq i, i'} \bar{G}_{i''} \left(\beta^i v^i / v^{i''} \right) g_{i'}(\beta^i v^i / v^{i'}) g_i(\beta^i) d\beta^i,$$

$$\frac{\partial^2 \Pi}{\partial v^i \partial v^{i'}} = \int_0^1 \frac{a(\beta^i)^2 v^i}{(hK\beta^i v^i + 1)^2 (v^{i'})^2} \prod_{i'' \neq i, i'} \bar{G}_{i''} \left(\beta^i v^i / v^{i''} \right) g_{i'}(\beta^i v^i / v^{i'}) g_i(\beta^i) d\beta^i. \tag{4.47}$$

We may not be able to calculate integrals in (4.11) and (4.47) analytically. Instead, we approximate them numerically using a Gaussian quadrature rule, in particular, the adaptive Gauss-Kronrod quadrature rule [119]. The idea behind the Gaussian quadrature rules is to factor the integrand into a nonnegative weight function and a function that is well approximated by a polynomial. Then, an N-point quadrature formula approximates the integral within a finite interval by the sum of the integrand's functional values at a set of N aptly chosen points (abscissas), multiplied by certain weighting coefficients. Note that the locations of the abscissas are not necessarily equally spaced. The interested reader is referred to Stoer and Bulirsch [174] for further details on the Gaussian quadrature rules.

4.3.2 Solving the Committee's Problem

We propose a branch-and-price algorithm to solve the CP. Observe that given the strain selection decisions $\widehat{\mathbf{X}}_\ell$ at node $\ell \in \mathcal{N}_{T-1}$, we can solve the MP to get expected flu shot production $\widehat{W}_\ell$, and calculate the expected societal vaccination benefit $\widehat{\Psi}_\ell(\delta_\ell, W_\ell)$.

For $\ell \in \mathcal{N}_{T-1}$, define binary variable $y^i_{m\ell} = 1$ iff strain $i \in \mathcal{I}$ is selected at node $m \in \mathscr{P}(\ell)$. Let $\mathbf{y}_{m\ell} = (y^i_{m\ell})_{i\in\mathcal{I}}$ and $\mathbf{x}_m = (x^i_m)_{i\in\mathcal{I}}$. The set of all feasible $(\mathbf{y}_{m\ell})_{m\in\mathscr{P}(\ell)}$ for node $\ell \in \mathcal{N}_{T-1}$ is given by

$$\mathcal{Y}_\ell = \left\{ y^i_{m\ell} \in \{0,1\},\ i \in \mathcal{I}, m \in \mathscr{P}(\ell) \;\middle|\; \sum_{m\in\mathscr{P}(\ell)} y^i_{m\ell} \le 1,\ i \in \mathcal{I} \right\}. \tag{4.48a}$$

Let $\mathcal{D}_\ell$ denote the index set of $\mathcal{Y}_\ell$, i.e. $\mathcal{Y}_\ell = \{(\hat{\mathbf{y}}_{m\ell})^\tau_{m\in\mathscr{P}(\ell)} \mid \tau \in \mathcal{D}_\ell\}$. Note that $\mathcal{D}_\ell$ is finite, and we can express any element of $\mathcal{Y}_\ell$ by

$$(\mathbf{y}_{m\ell})_{m\in\mathscr{P}(\ell)} = \sum_{\tau\in\mathcal{D}_\ell} \lambda^\tau_\ell (\hat{\mathbf{y}}_{m\ell})^\tau_{m\in\mathscr{P}(\ell)},\ \sum_{\tau\in\mathcal{D}_\ell} \lambda^\tau_\ell = 1,\ \lambda^\tau_\ell \in \{0,1\},\ \tau \in \mathcal{D}_\ell.$$

Each $(\hat{\mathbf{y}}_{m\ell})^\tau_{m\in\mathscr{P}(\ell)}$ corresponds to an expected societal vaccination benefit $\Psi_\ell(\hat{\delta}^\tau_\ell, \widehat{W}^\tau_\ell)$ that maximizes the objective function (3.2). We substitute $(\mathbf{y}_{m\ell})_{m\in\mathscr{P}(\ell)}$ $\forall \ell \in \mathcal{N}_{T-1}$ to obtain the Dantzig-Wolfe reformulation of the CP.

$$\text{[CP-DW]} \quad \max \sum_{\ell\in\mathcal{N}_{T-1}} p_\ell \left[\sum_{\tau\in\mathcal{D}_\ell} \Psi_\ell(\hat{\delta}^\tau_\ell, \widehat{W}^\tau_\ell)\lambda^\tau_\ell \right], \tag{4.49}$$

$$\text{subject to} \quad \sum_{\tau\in\mathcal{D}_\ell} \hat{\mathbf{y}}^\tau_{m\ell}\lambda^\tau_\ell = \mathbf{x}_m \qquad \ell \in \mathcal{N}_{T-1}, m \in \mathscr{P}(\ell), \tag{4.50}$$

$$\sum_{\tau\in\mathcal{D}_\ell} \lambda^\tau_\ell = 1 \qquad \ell \in \mathcal{N}_{T-1}, \tag{4.51}$$

$$\lambda^\tau_\ell \in \{0,1\},\ \mathbf{x}_m \ge 0 \qquad \forall m \in \mathcal{N}_{T-1}, \ell \in \mathscr{P}(m), \tau \in \mathcal{D}_\ell. \tag{4.52}$$

Constraints (4.50) ensure nonanticipativity. Convexity constraints (4.51) choose exactly one solution from set $\mathcal{D}_\ell$ for each node $\ell \in \mathcal{N}_{T-1}$. Note that CP-DW does not impose binary restrictions on $\mathbf{x}_m$ variables as they are satisfied for any binary vector λ.

The number of columns in DW-CP is exponential in that of CP, however, its LP relaxation is much stronger (see Section 4.5.2). We solve CP-DW using a branch-and-price algorithm. The main advantage of this approach is that it optimizes over the binary variables $\mathbf{X}_\ell,\ \ell \in$

$\mathcal{N}_{T-1}$, while computing the corresponding values of the continuous variables in the pricing subproblem.

The cardinality of $\mathcal{D}_\ell$ is huge, even for moderate-sized instances. We first create a restricted master problem (CP-RMP) in which each set $\mathcal{D}'_\ell$ represents a modest-sized subset of $\mathcal{D}_\ell$, $\ell \in \mathcal{N}_{T-1}$. Given the optimal duals, $\hat{\pi}_{m\ell}$ and $\hat{\mu}_\ell$ that correspond to constraints (4.50) and (4.51) in the LP relaxation of the CP-RMP, we identify the column $\tilde{\tau} \in \mathcal{D}_\ell$ that has the most favorable reduced cost by solving a pricing subproblem.

$$[\text{CP-SP}(\ell)] \quad \max \Psi_\ell(\delta_\ell, W_\ell) - \sum_{i\in\mathcal{I}} \sum_{m\in\mathscr{P}(\ell)} \hat{\pi}^i_{m\ell} y^i_{m\ell} - \hat{\mu}_\ell, \tag{4.53}$$

$$\text{subject to} \quad (\mathbf{y}_{m\ell})_{m\in\mathscr{P}(\ell)} \in \mathcal{Y}_\ell, \tag{4.54}$$

$$\mathbf{v}_{k\ell} \in \operatorname{argmax} \Pi_{k\ell} \qquad k \in \mathcal{K}, \tag{4.55}$$

$$W_\ell = \sum_{k\in\mathcal{K}} \mathbb{E}[\min_{i\in\mathcal{I}(\ell)} \{\beta^i_n v^i_{k\ell}\}], \tag{4.56}$$

$$\delta^{ij}_\ell \le M \min \left\{ 1 - \sum_{m\in\mathscr{P}(\ell)} y^h_{m\ell}, \sum_{m\in\mathscr{P}(\ell)} y^i_{m\ell} \right\} \qquad i,j \in \mathcal{I}, h \in \mathcal{H}(i,j), \tag{4.57}$$

$$\delta^{ij}_\ell \ge 0 \qquad i,j \in \mathcal{I}. \tag{4.58}$$

We solve CP-SP(ℓ) using dynamic programming (see the Appendix3.3.3) over the $\mathbf{y}_{m\ell}$ variables. This approach improves convergence of the column generation by considering all feasible columns generated at different iterations. The values of the continuous variables are calculated by solving the Manufacturer's Problem.

Any feasible solution to CP-SP(ℓ) with a positive objective value $z_{CP-SP}(\ell) > 0$ lets us create a new column for CP-RMP, i.e. add a new element to $\mathcal{D}'_\ell$. If no such solution exists for any $\ell \in \mathcal{N}_{T-1}$, then we have solved the LP relaxation of CP-DW (CP-DW-LP) optimally. If the optimal solution to CP-DW-LP is integral, then we have solved CP-DW. If not, the algorithm branches. For further technical details of the branch-and-price method for multi-stage stochastic integer programs, we refer the reader to [129].

4.4 PARAMETRIZATION

Our parameter values are based on the 2008-09 flu season in the United States. Parametrization of the flu infection cost (c), expected cost per unvaccinated person in the target population (q_n), cross-effectiveness among the candidate strains $b(i,j)$, and prevalence of strains (e_n^i) are the same as in Section 3.4.

Strain production yield ratios. We assume that the production yield ratios of strains (β^i) follow a Normal distribution that is truncated between 0.15 and 1, see Figure 4.1. We set the mean of low yield ratio to 0.7, the mean of moderate yield ratio to 0.8, and the mean of high yield ratio to 0.9. Moreover, we set the standard deviation of the production yield ratios to 0.2 for the current vaccine strains, and to 0.4 for the others.

Strain production cost. We assume that the variable production cost for all strains (φ^i) are equal. Moreover, to incorporate the effect of manufacturing difficulties due to decreased production time, we assume that the unit strain production cost is increasing in time. Deo and Corbett [56] and Cho [45] used the variable cost of \$3 per dose of the flu shot (with three strains). We assume that approximately 20% of this cost is due to filling and packaging operations after growing the strains. Therefore, we set the unit strain production cost to $\$(3-0.6)/3 = \$0.8$ in the first time stage. We increase this cost by \$0.2 in each time stage.

Target population for influenza vaccination. The 2008-2009 influenza season marked the first season where all healthy children ages 6 months up to 19 years old were recommended to get the vaccine [70]. With this expansion, 261 million people in the United States (D) were in the target population for influenza vaccination.

Price function parameters. We set the reservation price (a) to the cost of flu infection (c), which is \$41 per person. In the 2008-09 flu season, 136 million doses of the flu shot was brought to the market in the United States [90], and the average price per dose was \$11.8 [32]. Using those values in the fractional price function $\frac{a}{hQ+1}$, we solve for h, which is approximately 0.02 per million doses.

4.5 NUMERICAL EXPERIMENTS

4.5.1 Implementation

We implement our branch-and-price algorithm using BCP, a framework for branch, cut, and price algorithms [155]. The same adjustments to make the branch-and-price algorithm work efficiently are implemented as in Section 3.5.1.

We solve (4.11) using a modified version of Powell's Hybrid method [152] as implemented in [75]. We use the Gauss-Kronrod quadrature rule of the GNU Scientific Library [75] to approximate the integrals in (4.11) and (4.47). Note that the Manufacturers' Problem does not depend on the prevalence of strains. Therefore, if the same strains are selected under two different scenarios, then the solution of the Manufacturers' Problem for those two scenarios would be the same. We store the results of the Manufacturers' Problem obtained for different strain selection decisions, and use them under different scenarios.

4.5.2 Performance of the Proposed Solution Technique

We evaluate the performance of the proposed solution technique using 12 SSP instances whose characteristics are provided in Table 4.1. All computational experiments are conducted on an Intel Xeon PC with 3GHz CPU and 3GB of RAM using ten-hour time limit. Table 4.2 reports the performance of our branch-and-price algorithm. Note that we can not compare the solution time of our algorithm with off-the-shelve solvers, e.g. CPLEX 11.0 (2010), because the Committee's Problem (CP) must consider the equilibrium of the Manufacturers' Problem, which is obtained by solving a nonlinear system of equations (4.11). Our proposed branch-and-price algorithm solves all instances in less than two hours. Note that CP-DW-LP, the LP relaxation of the CP-DW at the root node of the branch-and-price tree, is often integer feasible.

4.5.3 Sensitivity Analysis

In this section, we examine how the level of yield uncertainty and the reservation price jointly affect the expected industry output and expected consumer surplus in equilibrium. We consider the M2 instance in Table 4.1 (i.e. 4 stages and moderate attack rate), and assume that the strain selection decisions are fixed to the optimum solution of this instance.

Each cell in Table 4.3 displays the ratio of expected industry output under uncertainty to that in the deterministic case. In each column, coefficient of variation $\delta^i = \mu^i/\sigma^i$ of the production yield of strains are multiplied by $\Delta\delta$ to increase the uncertainty. This table illustrates Proposition 4.6. In addition, note that the expected industry output is decreasing in uncertainty for fixed reservation price.

Each cell in Table 4.4 displays the ratio of expected consumer surplus under yield uncertainty to that of the deterministic case. In each column, coefficient of variation $\delta^i = \mu^i/\sigma^i$ of the production yield of strains are multiplied by $\Delta\varsigma$ to increase the uncertainty. This table verifies our observation in Section 4.2 that the expected consumer surplus under uncertainty exceeds that in the deterministic case above a threshold reservation price. In addition, note that the expected consumer surplus is decreasing in uncertainty for fixed reservation price.

Each cell in Table 4.5 displays the ratio of expected industry output under uncertainty to that in the deterministic case. This table is obtained after increasing the strain production cost by 50%. Recall that there are two basic factors that influence the optimal production quantity of a strain facing yield uncertainty. First, yield uncertainty increases the effective unit production cost, i.e. $\varphi^i/\beta^i \geq \varphi^i$, so we need to reduce the production quantity of that strain. Second, we need to inflate the production quantity of strain i to account for the underlying yield uncertainty. These two factors act against each other. When the flu shot price is low, the first factor is dominant and the manufacturer needs to decrease the production quantity. When the flu shot price is high, the second factor becomes dominant and the manufacturer needs to increase the production quantity. Table 4.5 verifies this observation as the threshold reservation prices where the expected industry output under uncertainty exceeds that in the deterministic case are higher than that in Table 4.3 in each column.

4.6 CONCLUSIONS

We formulate a two-level model for the annual influenza strain selection problem that incorporates both the society's and the vaccine manufacturers' perspectives. We analyze the effects of yield uncertainty, price, and cost parameters on manufacturers' production decisions. Our model considers the selection and production of multiple flu shot strains.

We find that if there is a single strain in the flu shot, then the expected industry output under yield uncertainty is always lower than that in the deterministic case. However, this is not true when there are multiple flu shot strains. Based on price of the vaccine, the expected industry output under yield uncertainty can exceed that in the deterministic case. The high flu shot price provides the manufacturer sufficient incentive to increase the production levels of strains in hope of a high level of realized yield.

Consumer surplus is increasing in the amount of the flu shot that is brought to the market. Therefore, if the manufacturers have enough incentive to increase production levels under uncertainty, the expected consumer surplus under yield uncertainty can exceed that in the deterministic case.

There are several cases that we have not explored here. For instance, the influenza vaccine manufacturers might have heterogenous and/or correlated yield distributions. Moreover, the manufacturers might have different production lead times, allowing the ones with smaller lead times to respond to yield realizations of the others with longer lead times. The hyperbolic demand function shapes some of our results; considering more general demand functions, although challenging, might be worthwhile.

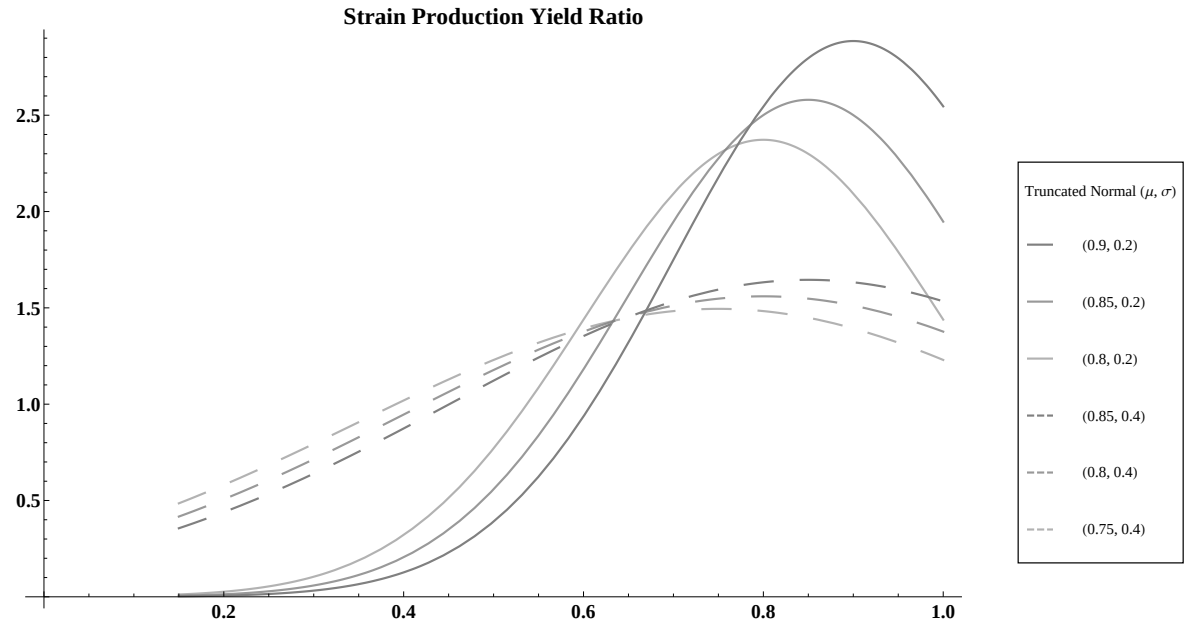


Figure 4.1: The mean of low yield ratio is 0.7, the mean of moderate yield ratio is 0.8, and the mean of high yield ratio is 0.9. Moreover, the standard deviation of the production yield ratio is 0.2 for the current vaccine strains, and 0.4 for the others. (i)A/NewCaledonia/20/99∼TrN(0.9,0.2),(ii)A/SolomonIslands/3/06∼TrN(0.8,0.4), (iii)A/Brisbane/59/07∼TrN(0.85,0.4),(iv)A/SouthDakota/6/07∼TrN(0.75,0.4), (v)A/Wisconsin/67/2005∼TrN(0.85,0.2),(vi)A/Brisbane/10/2007∼TrN(0.85,0.4), (vii)B/Malaysia/2506/2004∼TrN(0.8,0.2),(viii)B/Florida/04/2006∼TrN(0.75,0.4), (ix)B/Brisbane/03/2007∼TrN(0.8,0.4).

Table 4.1: Sizes of the test instances. The attack rate is the proportion of individuals exposed to an infectious agent who becomes clinically ill. The scenarios used under different attack rates are the same except the demand (d_n), shortage cost (q_n) and the prevalence (e_n^i) parameters as explained in Section 3.4. The number of weeks between two consecutive Committee meetings is denoted by Δt.

Attack		Δt	Scenario tree size			Extensive form size		
rate	Name	(weeks)	$\|T\|$	$\|\mathcal{N}_T\|$	$\|\mathcal{N}\|$	Cont. Var.	Bin. Var.	Constraints
Low	L1	4	3	100	111	1000	99	84800
	L2	2	4	512	585	5120	657	434176
	L3	1.5	5	243	341	2430	882	206064
	L4	1	6	32	63	320	279	27136
Moderate	M1	4	3	100	111	1000	99	84800
	M2	2	4	512	585	5120	657	434176
	M3	1.5	5	243	341	2430	882	206064
	M4	1	6	32	63	320	279	27136
High	H1	4	3	100	111	1000	99	84800
	H2	2	4	512	585	5120	657	434176
	H3	1.5	5	243	341	2430	882	206064
	H4	1	6	32	63	320	279	27136

Table 4.2: Performance of the proposed solution technique. DW-LP is the LP relaxation of the DW at the root node of the branch-and-price tree. % gap is calculated relative to the optimal solution.

	DW-LP	Branch-and-price	Optimal val.($M)
	% gap	sec.	CP
L1	0.00	8.13	571.66
L2	< 0.01	685.46	585.39
L3	1.13	5998.47	606.11
L4	0.00	268.91	591.99
M1	0.00	7.10	1155.95
M2	0.00	163.11	1162.60
M3	< 0.01	563.81	1205.42
M4	0.00	209.83	1159.22
H1	0.00	6.85	2435.59
H2	0.00	158.34	2482.38
H3	< 0.01	455.62	2563.72
H4	0.00	183.08	2495.15

Table 4.3: Impact of yield uncertainty on the expected industry output in equilibrium. Coefficient of variation $\delta^i = \mu^i/\sigma^i$ of the production yield of strains are multiplied by $\Delta\delta$ to increase the uncertainty

a	$\Delta\delta = 1.25$	$\Delta\delta = 1.50$	$\Delta\delta = 1.75$	$\Delta\delta = 2.00$	$\Delta\delta = 2.25$	$\Delta\delta = 2.50$	$\Delta\delta = 2.75$	$\Delta\delta = 3.00$
10	1.14	**0.94**	**0.80**	**0.70**	**0.62**	**0.56**	**0.51**	**0.47**
20	1.27	1.10	**0.98**	**0.90**	**0.83**	**0.78**	**0.74**	**0.70**
30	1.33	1.18	1.07	**0.99**	**0.92**	**0.87**	**0.83**	**0.79**
40	1.38	1.23	1.12	1.04	**0.98**	**0.93**	**0.89**	**0.85**
50	1.42	1.27	1.16	1.09	1.03	**0.98**	**0.93**	**0.89**
60	1.45	1.30	1.20	1.12	1.06	1.01	**0.97**	**0.93**
70	1.47	1.33	1.22	1.15	1.09	1.04	**0.99**	**0.96**
80	1.50	1.35	1.25	1.17	1.11	1.06	1.02	**0.98**
90	1.52	1.37	1.27	1.19	1.13	1.08	1.04	1.00
100	1.53	1.39	1.29	1.21	1.15	1.10	1.05	1.02

Table 4.4: Impact of yield uncertainty on the expected consumer surplus in equilibrium. Coefficient of variation $\delta^i = \mu^i/\sigma^i$ of the production yield of strains are multiplied by $\Delta\varsigma$ to increase the uncertainty.

a	$\Delta\delta = 1.25$	$\Delta\delta = 1.50$	$\Delta\delta = 1.75$	$\Delta\delta = 2.00$	$\Delta\delta = 2.25$	$\Delta\delta = 2.50$	$\Delta\delta = 2.75$	$\Delta\delta = 3.00$
10	1.20	**0.92**	**0.73**	**0.60**	**0.50**	**0.43**	**0.37**	**0.32**
20	1.29	1.11	**0.98**	**0.89**	**0.81**	**0.75**	**0.70**	**0.65**
30	1.32	1.17	1.07	**0.99**	**0.92**	**0.87**	**0.82**	**0.78**
40	1.33	1.20	1.11	1.04	**0.98**	**0.94**	**0.89**	**0.86**
50	1.33	1.22	1.14	1.07	1.02	**0.98**	**0.94**	**0.90**
60	1.34	1.23	1.16	1.10	1.05	1.01	**0.97**	**0.94**
70	1.34	1.24	1.17	1.11	1.07	1.03	**0.99**	**0.96**
80	1.34	1.25	1.18	1.13	1.08	1.05	1.01	**0.98**
90	1.34	1.25	1.19	1.14	1.10	1.06	1.03	1.00
100	1.34	1.26	1.19	1.15	1.11	1.07	1.04	1.01

Table 4.5: Impact of yield uncertainty on the expected industry output in equilibrium. Coefficient of variation $\delta^i = \mu^i/\sigma^i$ of the production yield of strains are multiplied by $\Delta\varsigma$ to increase the uncertainty. Strain production cost is increased by 50%.

a	$\Delta\delta = 1.25$	$\Delta\delta = 1.50$	$\Delta\delta = 1.75$	$\Delta\delta = 2.00$	$\Delta\delta = 2.25$	$\Delta\delta = 2.50$	$\Delta\delta = 2.75$	$\Delta\delta = 3.00$
10	1.06	**0.80**	**0.63**	**0.50**	**0.40**	**0.33**	**0.27**	**0.22**
20	1.20	1.01	**0.89**	**0.79**	**0.72**	**0.67**	**0.62**	**0.58**
30	1.27	1.10	**0.98**	**0.90**	**0.83**	**0.78**	**0.74**	**0.70**
40	1.31	1.16	1.05	**0.96**	**0.90**	**0.85**	**0.80**	**0.77**
50	1.35	1.20	1.09	1.01	**0.95**	**0.90**	**0.85**	**0.81**
60	1.38	1.23	1.12	1.04	**0.98**	**0.93**	**0.89**	**0.85**
70	1.41	1.26	1.15	1.07	1.01	**0.96**	**0.92**	**0.88**
80	1.43	1.28	1.18	1.10	1.04	**0.99**	**0.94**	**0.91**
90	1.45	1.30	1.20	1.12	1.06	1.01	**0.97**	**0.93**
100	1.47	1.32	1.22	1.14	1.08	1.03	**0.98**	**0.95**

5.0 BILEVEL CROSS-VALIDATION MODELS: AN APPLICATION TO INFLUENZA A/H3N2 VIRUS

Influenza viruses frequently mutate in response to antibody pressure. Those mutations gradually accumulate and cause antigenically distinct strains to circulate (i.e. antigenic drift). The annual influenza vaccine provides protection if its composition is antigenically similar to the circulating strains. Therefore, the vaccine strains have to be updated frequently based on clinical, virological and immunological influenza surveillance. In the virological surveillance, hemagglutinin inhibition (HI) assays are used to identify antigenic properties of the influenza viruses (see Table 5.1). However, this serology assay is labor-intensive and time-consuming [96]. As an alternative, representing the difference between the antigenic properties of the influenza strains using a distance (i.e. antigenic distance) can be a rapid indicator of the likelihood that the current vaccine will protect against a recently emerged strain, and also facilitates the study of the virus' evolution.

Antigenic properties of the influenza A/H3N2 virus are characterized by its hemagglutinin surface protein (HA) that consists of three identical subunits. Each subunit has two chains, HA1 and HA2, which are 329 and 175 amino acid (AA) residues long, respectively. We focus on the HA1 chain as it mutates more frequently than HA2 does [126]. Hence, by AA sequence of an influenza A/H3N2 virus, we refer to that of its HA1 chain. Smith et al. [171] and Lee and Chen [123] discuss the suitability of using the HA1 chain for predicting the antigenic variants of the influenza A/H3N2 virus.

Empirical studies have documented that one to three amino acid changes on the HA1 reduce the antigenicity and efficacy of inactivated vaccines [97, 101, 107, 143, 187], which indicate the existence of *immunodominant positions* in the HA1 [124]. Bush et al. [24] analyze 375 A/H3N2 viruses isolated between 1984 and 1996. Of the 329 AA positions on

Table 5.1: The FDA uses Hemagglutinin Inhibition (HI) assays to evaluate the cross-effectiveness among the candidate strains. Each cell shows the minimum antiserum concentration, raised against the reference influenza strain, that completely inhibits the agglutination of the test strain [67].

	Reference Strain (Antibody)			
Test Strain (Antigen)	NC/20	SI/3	AS/59	SD/6
A/New Caledonia/20/99	640	160	160	160
A/Solomon Islands/3/06	160	640	320	160
A/Brisbane/59/07	80	160	640	640
A/South Dakota/6/07	80	160	640	640

the HA1, eighteen AA positions are found to be positively associated with antibody response difference. In another study, Smith et al. [171] have identified twenty amino acid positions on the HA1 as antibody-binding sites. Lee and Chen [123] have compared the usefulness of those AA positions reported in Bush et al. [24] and Smith et al. [171] on predicting antigenic variants of influenza A/H3N2 viruses. Five regions of the HA1 chain have been identified to undergo mutations, eventually leading to antibody escape. A model based on the number of AA changes in those antigenic sites is found to perform well for predicting antigenic variety.

Liao et al. [126], Lee and Chen [123] and Wilson and Cox [186] propose statistical methods based on Hamming distance to predict antigenic variety. The pairwise comparisons of AA sequences of each strain are considered as independent variables and the corresponding antigenic distance between each pair are considered as dependent variables. As a result, AA positions that are positively associated with antigenic distances are identified as potential immunodominant positions that cause antigenic variety.

Smith et al. [171] provide an antigenic map of 253 A/H3N2 viruses isolated between 1968 and 2003, and cluster them into 11 antigenic groups (see Figure 5.1). Retrospective quantitative analysis of the genetic data have revealed that 45 AA positions are related to

cluster-difference substitutions. Huang et al. [95] use the "information gain" and "entropy" measures from machine learning [154] to rank amino acid positions for differentiating antigenic variants and similar viruses. An amino acid position with high information gain implies that this position is highly correlated to the antigenic variants. An amino acid position with high entropy means that this position is often mutated in the data set.

Machine learning methods train a linear regression model based on available data. They take a set of hyper-parameters including the features or variables that should be used in the model. The most commonly used and widely accepted method for selecting these hyper-parameters is cross validation [114]. In cross validation, hyper-parameters are selected to minimize some estimate of the out-of-sample generalization error. A typical method would define a grid over the hyper-parameters of interest, and then do ten-fold cross validation for each of the grid values. Note that the combinatorial explosion of grid points in high dimensions effectively limits the desirable number of hyper-parameters.

In this chapter, we formulate cross validation for selecting the variables of a linear regression model as a bilevel program in which an upper-level program chooses the model variables to minimize the out-of-sample error, and lower-level problems in each fold optimize in-sample errors according to their training data set by selecting the regression coefficients of the chosen model variables. This approach allows for optimizing the cross validation outcome. Moreover, it offers modeling flexibility when considering multiple statistical learning goals. In our numerical experiments, we use a serologic data set to identify AA positions on the HA1 chain which cause antigenic variety.

5.1 BILEVEL MODEL FORMULATION

Let $\mathcal{I} = \{1, \ldots, I\}$ be a finite set of A/H3N2 strains. Each strain $i \in \mathcal{I}$ is represented by a sequence of M amino acids whose index set $\{1, \ldots, M\}$ is denoted by $\mathcal{M}$. Let $\mathbf{x}_{ij} \in \{0,1\}^M$ be a binary vector obtained by pairwise comparison of the AA sequences of strains i and j. Denote by c_{ij} the HI assay value for those two strains. Then, the antigenic distance between

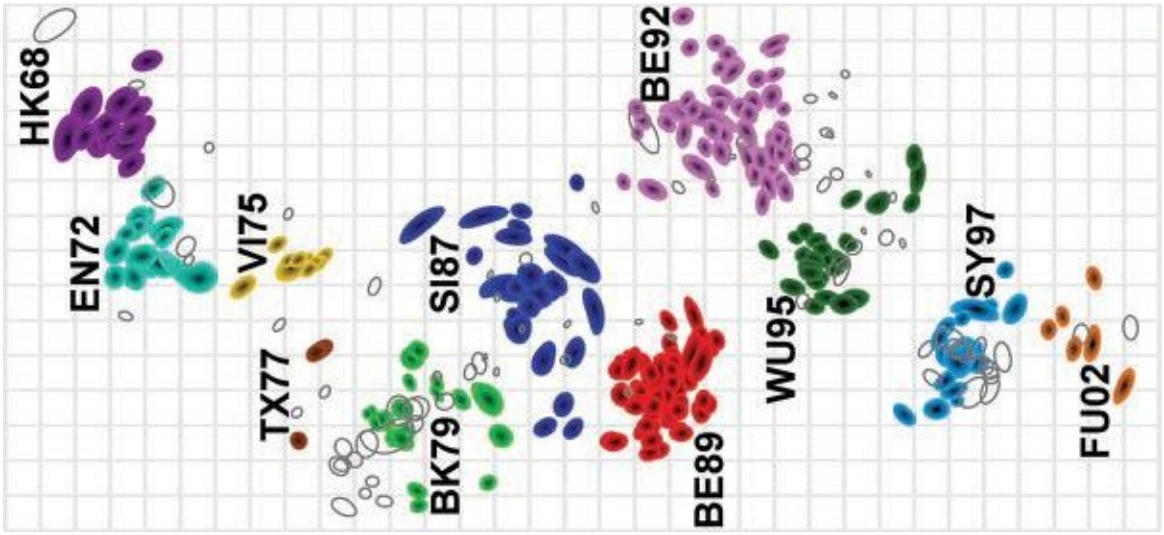


Figure 5.1: Antigenic map of influenza A /H3N2 virus from 1968 to 2003. Strain color represents the antigenic cluster to which the strain belongs. Clusters were identified by a k-means clustering algorithm and named after the first vaccine strain in the clustertwo letters refer to the location of isolation and two digits refer to year of isolation. The vertical and horizontal axes both represent antigenic distance, and, because only the relative positions of antigens and antisera can be determined, the orientation of the map within these axes is free [171].

i and j is given by $y_{ij} = \left(\sqrt{c_{ij}c_{ji}/c_{ii}c_{jj}}\right)^{-1}$ [105]. Two strains are considered as *antigenic variants* if $y_{ij} \geq 4$ [123]. We refer to this cutoff value as the *classifier margin*.

Let $\Omega = \{(\mathbf{x}_{ij}, y_{ij})\}_{\forall i,j} \subset \{0,1\}^M \times \mathbb{R}^1$. For T-fold cross validation, Ω is partitioned into T disjoint validation sets Ω^t, $t = 1, \ldots, T$. Denote by $\mathcal{T}$ the index set of folds $\{1, \ldots, T\}$. Subsequent T iterations are performed such that within each iteration a different fold of the data is designated for validation while the remaining $T-1$ folds are used for training. Let $\bar{\Omega}^t = \Omega \setminus \Omega^t$ be the training set of fold $t \in \mathcal{T}$. We denote the vector of descriptors trained for fold t's inner-level problem using $\bar{\Omega}^t$ by $\pi^t \in \mathbb{R}^m$.

The objective function of the lower-level classification problem could be rather general. However, convex functions are often used so that the only non-convexity is generated by the optimality conditions [114]. We focus on the mean absolute deviation objective and formulate the T-fold bilevel multiple regression problem as:

$$[\text{AD}] \quad \min_{\mathbf{u}\in\{0,1\}^M} \quad \frac{1}{T}\sum_{t=1}^{T}\frac{1}{|\Omega^t|}\sum_{(i,j)\in\Omega^t}\Big|\sum_{m\in\mathcal{M}} x_{ijm}\pi_m^t - y_{ij}\Big| \tag{5.1}$$

$$\text{subject to} \quad \sum_{m\in\mathcal{M}} u_m \leq U, \tag{5.2}$$

$$\pi^t \in \operatorname{argmin}_{0\leq\pi\leq\Pi\mathbf{u}} \quad \sum_{(i,j)\in\bar{\Omega}^t}\Big|\sum_{m\in\mathcal{M}} x_{ijm}\pi_m - y_{ij}\Big| \quad \forall t \in \mathcal{T}. \tag{5.3}$$

If binary variable $u_m = 1$, then AA position m is included in the model and it can have a positive descriptor in the lower-level. Parameter Π is an upper bound on the antigenic distance and parameter U is the maximum number of AA positions that can be included in the model. There are T lower-level problems (5.3) that model the training of descriptors π^t within each fold. The upper-level objective (5.1) minimizes the mean absolute deviation over all folds. We discuss two more upper-level objective functions.

The misclassification objective (MC): Let parameter $h_{ij} = 1$ if i and j are antigenic variants ($y_{ij} \geq 4$), and $h_{ij} = 0$ otherwise. Strains i and j are *misclassified* in fold t, if they are antigenic variants but the regression model of that fold predicts them as antigenically similar, or vice versa. We define binary variable $z_{ij}^t = 1,\ (i,j) \in \Omega^t$ iff i and j are misclassified in fold t, and introduce two additional constraints to the upper-level problem.

$$\sum_{m \in \mathcal{M}} x_{ijm}\pi_m^t \geq 4h_{ij}(1 - z_{ij}^t) \qquad t \in \mathcal{T}, (i,j) \in \Omega^t, \tag{5.4}$$

$$(1 - h_{ij}) \sum_{m \in \mathcal{M}} x_{ijm}\pi_m^t \leq 4 + \Pi z_{ij}^t \qquad t \in \mathcal{T}, (i,j) \in \Omega^t. \tag{5.5}$$

Constraints (5.4) and (5.5) are inactive when $z_{ij}^t = 1$. Moreover, if $h_{ij} = 1$, constraint (5.5) is inactive and constraint (5.4) is satisfied with $\sum_{m \in \mathcal{M}} x_{ijm}\pi_m^t \geq 4$ when $z_{ij} = 0$. Likewise, if $h_{ij} = 0$, then constraint (5.4) is inactive and constraint (5.5) is satisfied with $\sum_{m \in \mathcal{M}} x_{ijm}\pi_m^t \leq 4$ when $z_{ij} = 0$. Then we formulate the objective function that minimizes the misclassification of the validation points as:

$$\text{[MC]} \quad \min_{\mathbf{u},\mathbf{z}} \quad \frac{1}{T}\sum_{t=1}^{T} \frac{1}{|\Omega^t|} \sum_{(i,j) \in \Omega^t} z_{ij}^t. \tag{5.6}$$

The hinge loss objective (HL): For each fold t, the *hinge loss* is the average distance of misclassified validation points from the classifier margin. Let h_{ij} and z_{ij}^t be as defined for the MC. We formulate the hinge loss minimization objective for the upper-level problem as:

$$\text{[HL]} \quad \min_{\mathbf{u},\mathbf{z}} \quad \frac{1}{T}\sum_{t=1}^{T} \frac{1}{|\Omega^t|} \sum_{(i,j) \in \Omega^t} z_{ij}^t |4 - \sum_{m \in \mathcal{M}} x_{ijm}\pi_m^t|. \tag{5.7}$$

To linearize the HL, we define variable $\delta_{ij}^t \geq 0$, and introduce the following constraint to the upper-level problem in addition to the constraints (5.4)-(5.5).

$$\delta_{ij}^t \geq (1 - 2h_{ij})(\sum_{m \in \mathcal{M}} x_{ijm}\pi_m^t - 4) - \Pi(1 - z_{ij}^t) \qquad t \in \mathcal{T}, (i,j) \in \Omega^t. \tag{5.8}$$

Note that constraint (5.8) is inactive when $z_{ij}^t = 0$. Consider the case when $z_{ij}^t = 1$. If $h_{ij} = 1$, then constraint (5.8) is satisfied with $\delta_{ij}^t \geq 4 - \sum_{m \in \mathcal{M}} x_{ijm}\pi_m^t$. Likewise, if $h_{ij} = 0$, then constraint (5.8) is satisfied with $\delta_{ij}^t \geq \sum_{m \in \mathcal{M}} x_{ijm}\pi_m^t - 4$. Then we reformulate the HL as:

$$\text{[HL]} \quad \min_{\mathbf{u},\mathbf{z},\delta} \quad \frac{1}{T}\sum_{t=1}^{T} \frac{1}{|\Omega^t|} \sum_{(i,j) \in \Omega^t} \delta_{ij}^t. \tag{5.9}$$

5.2 SOLUTION APPROACH

We convert the AD into a mathematical program with linear equilibrium constraints. For fixed upper-level decision $\hat{\mathbf{u}} \in \{0,1\}^M$, the AD decomposes into T independent linear programs, one for each fold $t \in \mathcal{T}$.

$$[\text{LW}^t] \quad \min_{\pi} \quad \sum_{(i,j)\in\bar{\Omega}^t} \Big| \sum_{m\in\mathcal{M}} x_{ijm}\pi_m - y_{ij} \Big| \tag{5.10}$$

$$\text{subject to} \quad 0 \le \pi_m \le \Pi\hat{u}_m \qquad m \in \mathcal{M}. \tag{5.11}$$

An equivalent reformulation of the LW^t is given by,

$$[\text{RLW}^t] \quad \min_{\pi,f} \quad \sum_{(i,j)\in\bar{\Omega}^t} f_{ij}^t \tag{5.12}$$

$$\text{subject to} \quad 0 \le \pi_m \le \Pi\hat{u}_m \qquad m \in \mathcal{M}, \tag{5.13}$$

$$f_{ij}^t \ge \mathbf{x}_{ij}'\pi - y_{ij} \qquad (i,j) \in \bar{\Omega}^t, \tag{5.14}$$

$$f_{ij}^t \ge y_{ij} - \mathbf{x}_{ij}'\pi \qquad (i,j) \in \bar{\Omega}^t, \tag{5.15}$$

where auxiliary variables f_{ij}^t $(i,j) \in \bar{\Omega}^t$ and constraints (5.14)-(5.15) are introduced to eliminate the nonlinear absolute value term in (5.10). Let α^t and $\beta_{ij}^t, \gamma_{ij}^t$ be the dual variables associated with the constraints (5.13)-(5.15), respectively. Then, the dual of the RLW^t is given by,

$$[\text{D-RLW}^t] \quad \max \quad -\Pi \sum_{m\in\mathcal{M}} \hat{u}_m\alpha_m^t + \sum_{(i,j)\in\bar{\Omega}^t} y_{ij}\left(\gamma_{ij}^t - \beta_{ij}^t\right) \tag{5.16}$$

$$\text{subject to} \quad -\alpha_m^t + \sum_{(i,j)\in\bar{\Omega}^t} x_{ijm}(\gamma_{ij}^t - \beta_{ij}^t) \le 0 \qquad m \in \mathcal{M}, \tag{5.17}$$

$$\beta_{ij}^t + \gamma_{ij}^t = 1 \qquad (i,j) \in \bar{\Omega}^t, \tag{5.18}$$

$$\alpha_m^t, \beta_{ij}^t, \gamma_{ij}^t \ge 0 \qquad m \in \mathcal{M}, (i,j) \in \bar{\Omega}^t. \tag{5.19}$$

From strong duality, we can reformulate the AD as,

$$\text{[RAD]} \quad \min \quad \frac{1}{T}\sum_{t\in\mathcal{T}}\frac{1}{|\Omega^t|}\sum_{(i,j)\in\Omega^t} f_{ij}^t \tag{5.20}$$

$$\text{subject to} \quad 0 \le \pi_m^t \le \Pi u_m \qquad m\in\mathcal{M}, t\in\mathcal{T}, \tag{5.21}$$

$$\sum_{m\in\mathcal{M}} u_m \le U, \tag{5.22}$$

$$\sum_{(i,j)\in\bar{\Omega}^t} f_{ij}^t = -\Pi\sum_{m\in\mathcal{M}} u_m\alpha_m^t + \sum_{(i,j)\in\bar{\Omega}^t} y_{ij}\left(\gamma_{ij}^t - \beta_{ij}^t\right) \; t\in\mathcal{T}, \tag{5.23}$$

$$f_{ij}^t \ge \mathbf{x}_{ij}'\pi^t - y_{ij}, \quad f_{ij}^t \ge y_{ij} - \mathbf{x}_{ij}'\pi^t \qquad t\in\mathcal{T}, (i,j)\in\Omega, \tag{5.24}$$

$$-\alpha_m^t + \sum_{(i,j)\in\bar{\Omega}^t} x_{ijm}(\gamma_{ij}^t - \beta_{ij}^t) \le 0 \qquad m\in\mathcal{M}, t\in\mathcal{T}, \tag{5.25}$$

$$\beta_{ij}^t + \gamma_{ij}^t = 1 \qquad t\in\mathcal{T}, (i,j)\in\bar{\Omega}^t, \tag{5.26}$$

$$u_m \in \{0,1\}, \alpha_m^t, \beta_{ij}^t, \gamma_{ij}^t \ge 0 \qquad m\in\mathcal{M}, t\in\mathcal{T}, (i,j)\in\bar{\Omega}^t. \tag{5.27}$$

In the RAD, we replace the lower-level optimality constraint (5.3) by dual feasibility (5.25)-(5.26), primal feasibility (5.24) and the strong duality condition for the lower-level problems (5.23). Note that there exists a nonlinear expression $u_m\alpha_m^t$ in constraints (5.23).

Proposition 5.1. *In the optimal solution of the RAD* $\alpha_m^{t*} \le \sum_{(i,j)\in\bar{\Omega}^t} x_{ijm}^*$ *for all* $m \in \mathcal{M},\ t\in\mathcal{T}$.

Proof. Constraint (5.17) is satisfied at equality in the optimal solution of the D-RLW^t, because otherwise the objective value can be increased by decreasing the α_m^t that satisfies $\alpha_m^t > \sum_{(i,j)\in\bar{\Omega}^t} x_{ijm}(\gamma_{ij}^t - \beta_{ij}^t)$. It follows from (5.25) and (5.26) that

$$\alpha_m^{t*} = \sum_{(i,j)\in\bar{\Omega}^t} x_{ijm}(2\gamma_{ij}^{t*} - 1) \le \sum_{(i,j)\in\bar{\Omega}^t} x_{ijm}^*.$$

□

From Proposition 5.1 and the fact that $u_m \in \{0,1\}$, we can linearize $u_m \alpha_m^t$ by using an auxiliary variable $z_m^t \geq 0,\ m \in \mathcal{M},\ t \in \mathcal{T}$.

$$
\begin{aligned}
z_m^t &\leq u_m \sum_{(i,j)\in\bar{\Omega}^t} x_{ijm}, \\
z_m^t &\geq \alpha_m^t + (u_m - 1) \sum_{(i,j)\in\bar{\Omega}^t} x_{ijm}, \\
z_m^t &\leq \alpha_m^t + (1 - u_m) \sum_{(i,j)\in\bar{\Omega}^t} x_{ijm}.
\end{aligned}
$$

As a result, we obtain a linear mixed 0-1 programming problem. Our reformulation and linearization technique requires $M \times T$ auxiliary continuous variables. In the literature, bilevel programs with LPs in the lower-level are usually reformulated as MIPs by using complementary slackness rather than strong duality [11]. In our case, this approach would require $2(M \times T + |\Omega|)$ auxiliary binary variables.

5.3 NUMERICAL EXPERIMENTS

The simplest AA sequence alignment method assigns zero or one to each position based on whether two AA residues at that position are identical or not (i.e. the Hamming distance [80]). More advanced alignment methods incorporate polarity, charge and structure of AAs by grouping them into similarity classes [64]. Those methods assign zero to each position in the sequence if the two AA residues at that position belong to the same group, or one otherwise. As can be seen in Table 5.2, we use 6 different grouping methods introduced by Liao et al. [126].

Our data set is extracted from the online supplement of Liao et al. [126]. It contains 277 AA pairwise sequence alignment of 62 different AAs and the antigenic distance for each of the those pairs. We combine the RAD with different grouping methods and upper-level objective functions to identify the potential immunodominant AA positions. Note that the RAD returns a different regression model for each fold $t \in \mathcal{T}$. However, the same AA positions are used in each model. We set the maximum number of AA positions that can be included in

Table 5.2: Similarity classes of grouping methods for amino acid sequence alignment [126].

GM1	{non-polar: A, F, G, I, L, M, P, V, W}, {polar: C, N, Q, S, T, Y}, {charged: D, E, H, K, R}
GM2	{non-polar aliphatic: A, G, I, L, M, V}, {non-polar aromatic: F, P, W}, {polar: C, N, Q, S, T, Y}, {charged: D, E, H, K, R}
GM3	{non-polar: A, F, G, I, L, M, P, V, W}, {polar: C, N, Q, S, T, Y}, {positively charged: H, K, R}, {negatively charged: D, E}
GM4	{non-polar aliphatic: A, G, I, L, M, V}, {non-polar aromatic: F, P, W}, {polar: C, N, Q, S, T, Y}, {positively charged: H, K, R}, {negatively charged: D, E}
GM5	{non-polar aliphatic: A, I, L, M, P, V}, {non-polar aromatic: F, W, Y}, {polar: N, Q, S, T}, {positively charged: H, K, R}, {negatively charged: D, E}, {C}, {G}
GM6	{non-polar aliphatic: A, I, L, M, P, V}, {non-polar aromatic: F, W, Y}, {polar: N, Q, S, T}, {charged: D, E, H, K, R}, {C}, {G}

the model, $U = 5, 10, 15, 20$ and 25. For each value of U, we first solve the RAD with CPLEX 11.0 [48] using 1-hour time limit. After identifying the immunodominant AA positions, we build a regression model based on the first 181 data points and use the remaining 96 data points for validation. We compute the agreement, sensitivity and specificity for each U and grouping method using the AD, MC, and HL objective functions. The agreement rate is defined as the ratio of all truly predicted pairs to the number of all strain pairs. The ratio of predicted variants to true variants and the ratio of predicted similar viruses to true similar viruses are defined as sensitivity and specificity, respectively. We report the results of this experiment in Tables 5.3, 5.5 and 5.7.

The highest agreement rate in the validation set is 0.938, which is achieved by the MC model using 5 variables and the grouping methods 1 and 3. The highest agreement rate in the training data set is 0.896, which is achieved by the HL model using 25 variables and the grouping method 5. The agreement rate of the "no grouping" method is smaller than that of any grouping method when there are more than 5 variables in the model. The regression models considered here, in common with reported by many others (Lee and Chen [123], Liao et al. [126], Huang et al. [95]), correctly identify the vast majority of antigenic variants (high sensitivity), but tend to over predict the number of antigenic variants (relatively lower specificity). On examination of the results, we observe that a high number of false positives have relatively few mutations in the regions under consideration. To create a model with more balanced predictive properties, we can introduce a screening step, in which all positive results are examined, and those which has fewer than a specified cutoff number of AA changes in the regions under consideration are rejected.

Smith et al. [171] provide an antigenic map of 253 A/H3N2 strains (see Figure 5.1). They define 11 clusters based on 31878 pair-wise AA sequence comparisons. We use the regression models in Tables 5.3, 5.5 and 5.7 on the same data set. We assume that a pair of viruses are antigenic variants if they are from different clusters, or similar if they are from the same cluster. Smith et al. [171] report 45 AA positions associated with the cluster-difference substitutions. As seen in Tables 5.4, 5.6 and 5.8, our approach reached over 90% agreement rate by using only five AA positions.

The highly consistent result reveals that our model indeed reduces the positions related to cluster transitions to a small number and could still stand comparison with the antigenic map. However, it is hard to compare the performances of the two studies because Smith et al. [171] use a clustering algorithm to determine the clusters in the antigenic map, which are not always consistent with antigenic distances. Ideally, the real performances should be compared by using the same serological data. Furthermore, we could not solve the MC and the HL instances optimally in 1-hour time limit using CPLEX 11.0 [48]. Therefore, the reported statistics are based on the best available solution.

5.4 CONCLUSIONS

Our approach allows for optimizing the cross validation outcome when selecting model parameters for identifying antigenic variety among the influenza viruses. Such tools could be readily integrated to the global influenza surveillance system for rapid identification of the circulating strains that are antigenically different from the current vaccine strains. Our numerical results are inconclusive about the best model for predicting the antigenic variants. However, over 90% agreement rate is reached by using only 5 AA positions in the regression models.

Our objective functions either minimize the deviation from a given target (i.e. the AD, HL) or minimize the misclassification (i.e. the MC), which is equivalent to maximizing the agreement rate. As can be seen in Tables 5.3, 5.5 and 5.7, optimizing the agreement rate caused unbalanced predictive power between sensitivity and specificity. As a remedy, we can control the ratio of sensitivity to specificity by using a weighted sum of those two statistics in the upper-level objective function.

Finally, we use the absolute deviation objective in the lower-level objective for training the regression models in each fold. This enables us to reformulate the lower-level problems as a linear program. The benefit of using other objective functions such as the sum of squared deviations is an open question, however non-convex functions require specialized solution techniques as they would prohibit the reformulation of the bilevel model as an MIP.

Table 5.3: Agreement, sensitivity, and specificity of the AD model. The instances are solved with CPLEX 11.0 using a one-hour time limit.

Num. of	Grouping	Validation Set			Training Set			MIP gap%
Variables	Method	Agreement	Sensitivity	Specificity	Agreement	Sensitivity	Specificity	CPLEX
5	No Grouping	0.823	1.000	0.785	0.782	0.830	0.697	0.0
	GM1	0.880	0.882	0.880	0.818	0.796	0.856	0.0
	GM2	0.880	0.882	0.880	0.815	0.796	0.848	0.0
	GM3	0.880	0.882	0.880	0.812	0.800	0.833	0.0
	GM4	0.766	1.000	0.715	0.809	0.809	0.811	0.0
	GM5	0.880	0.882	0.880	0.851	0.848	0.856	0.0
	GM6	0.870	0.882	0.867	0.818	0.857	0.750	0.0
10	No Grouping	0.594	1.000	0.506	0.837	0.965	0.614	8.6
	GM1	0.823	1.000	0.785	0.876	0.887	0.856	0.0
	GM2	0.813	1.000	0.772	0.865	0.904	0.795	2.8
	GM3	0.750	1.000	0.696	0.809	0.870	0.705	1.2
	GM4	0.823	1.000	0.785	0.796	0.852	0.697	3.7
	GM5	0.823	1.000	0.785	0.829	0.926	0.659	6.6
	GM6	0.813	1.000	0.772	0.831	0.913	0.689	7.4
15	No Grouping	0.654	0.953	0.590	0.850	0.941	0.691	15.5
	GM1	0.810	1.000	0.770	0.862	0.934	0.736	8.4
	GM2	0.806	1.000	0.765	0.851	0.932	0.709	8.8
	GM3	0.823	1.000	0.785	0.851	0.922	0.727	9.2
	GM4	0.790	1.000	0.744	0.852	0.927	0.721	10.6
	GM5	0.819	1.000	0.780	0.842	0.920	0.706	13.1
	GM6	0.815	1.000	0.775	0.856	0.910	0.764	13.1
20	No Grouping	0.610	1.000	0.527	0.860	0.953	0.697	16.1
	GM1	0.806	1.000	0.765	0.872	0.923	0.782	9.5
	GM2	0.806	1.000	0.765	0.859	0.932	0.730	10.4
	GM3	0.810	1.000	0.770	0.859	0.927	0.739	11.3
	GM4	0.806	1.000	0.765	0.866	0.937	0.742	10.5
	GM5	0.796	0.871	0.780	0.846	0.934	0.694	13.0
	GM6	0.767	1.000	0.716	0.856	0.937	0.715	13.0
25	No Grouping	0.644	0.929	0.582	0.857	0.939	0.715	16.0
	GM1	0.781	1.000	0.734	0.861	0.930	0.739	9.5
	GM2	0.777	1.000	0.729	0.869	0.920	0.779	10.1
	GM3	0.769	0.906	0.739	0.870	0.934	0.758	10.0
	GM4	0.806	1.000	0.765	0.859	0.936	0.724	11.0
	GM5	0.806	1.000	0.765	0.863	0.943	0.724	13.4
	GM6	0.763	1.000	0.711	0.859	0.939	0.718	12.6

Table 5.4: Agreement, sensitivity, and specificity of the AD model on the data set of Smith et al. [171].

Num. of Variables	Grouping Method	Smith Set		
		Agreement	Sensitivity	Specificity
5	No Grouping	0.912	0.974	0.563
	GM1	0.928	0.978	0.641
	GM2	0.928	0.987	0.594
	GM3	0.927	0.985	0.602
	GM4	0.927	0.984	0.601
	GM5	0.924	0.978	0.619
	GM6	0.926	0.966	0.699
10	No Grouping	0.908	0.997	0.406
	GM1	0.957	0.997	0.734
	GM2	0.949	0.996	0.680
	GM3	0.921	0.990	0.528
	GM4	0.924	0.996	0.516
	GM5	0.907	0.994	0.411
	GM6	0.933	0.996	0.576
15	No Grouping	0.921	0.996	0.501
	GM1	0.918	0.998	0.463
	GM2	0.925	0.998	0.511
	GM3	0.927	0.995	0.545
	GM4	0.917	0.996	0.467
	GM5	0.916	0.994	0.478
	GM6	0.938	0.997	0.607
20	No Grouping	0.923	0.999	0.493
	GM1	0.937	0.998	0.589
	GM2	0.923	0.998	0.498
	GM3	0.921	0.998	0.485
	GM4	0.922	0.997	0.496
	GM5	0.925	0.996	0.519
	GM6	0.920	0.998	0.481
25	No Grouping	0.926	0.997	0.525
	GM1	0.919	0.999	0.466
	GM2	0.930	0.998	0.542
	GM3	0.919	0.997	0.475
	GM4	0.915	0.998	0.446
	GM5	0.921	0.997	0.492
	GM6	0.919	0.998	0.471

Table 5.5: Agreement, sensitivity, and specificity of the MC model. The instances are solved with CPLEX 11.0 using a one-hour time limit.

Num. of	Grouping	Validation Set			Training Set			MIP gap%
Variables	Method	Agreement	Sensitivity	Specificity	Agreement	Sensitivity	Specificity	CPLEX
5	No Grouping	0.688	1.000	0.620	0.834	0.909	0.705	46.8
	GM1	0.917	0.824	0.937	0.812	0.778	0.871	15.5
	GM2	0.938	0.765	0.975	0.796	0.730	0.909	20.9
	GM3	0.938	0.765	0.975	0.776	0.713	0.886	38.1
	GM4	0.880	0.882	0.880	0.804	0.791	0.826	45.2
	GM5	0.703	1.000	0.639	0.881	0.917	0.818	46.9
	GM6	0.823	1.000	0.785	0.854	0.857	0.848	20.0
10	No Grouping	0.719	1.000	0.658	0.834	0.926	0.674	92.0
	GM1	0.740	0.941	0.696	0.845	0.874	0.795	82.9
	GM2	0.849	0.882	0.842	0.826	0.843	0.795	88.5
	GM3	0.823	1.000	0.785	0.884	0.922	0.818	87.9
	GM4	0.745	1.000	0.690	0.865	0.870	0.856	91.2
	GM5	0.880	0.882	0.880	0.831	0.874	0.758	88.4
	GM6	0.823	1.000	0.785	0.892	0.917	0.848	86.4
15	No Grouping	0.615	1.000	0.532	0.850	0.922	0.724	96.4
	GM1	0.817	0.976	0.782	0.873	0.922	0.788	93.0
	GM2	0.817	0.976	0.782	0.873	0.906	0.815	93.3
	GM3	0.823	1.000	0.785	0.888	0.936	0.806	93.6
	GM4	0.817	0.953	0.787	0.860	0.936	0.727	94.5
	GM5	0.779	1.000	0.732	0.881	0.937	0.782	93.0
	GM6	0.838	0.953	0.813	0.877	0.925	0.794	93.2
20	No Grouping	0.671	0.800	0.643	0.851	0.890	0.782	96.9
	GM1	0.746	1.000	0.691	0.886	0.934	0.803	92.7
	GM2	0.798	0.847	0.787	0.877	0.923	0.797	92.7
	GM3	0.744	1.000	0.689	0.886	0.922	0.824	93.6
	GM4	0.740	0.824	0.722	0.845	0.877	0.791	95.5
	GM5	0.835	0.906	0.820	0.873	0.925	0.782	93.8
	GM6	0.798	1.000	0.754	0.885	0.922	0.821	93.2
25	No Grouping	0.675	0.953	0.615	0.843	0.934	0.685	96.2
	GM1	0.725	0.776	0.714	0.817	0.845	0.767	94.3
	GM2	0.698	0.729	0.691	0.829	0.814	0.855	95.4
	GM3	0.760	0.894	0.732	0.850	0.885	0.788	95.5
	GM4	0.748	1.000	0.694	0.892	0.957	0.779	94.2
	GM5	0.792	1.000	0.747	0.881	0.946	0.767	94.0
	GM6	0.852	0.929	0.835	0.861	0.934	0.733	94.5

Table 5.6: Agreement, sensitivity, and specificity of the MC model on the data set of Smith et al. [171].

Num. of Variables	Grouping Method	Smith Set		
		Agreement	Sensitivity	Specificity
5	No Grouping	0.911	0.961	0.627
	GM1	0.894	0.944	0.612
	GM2	0.885	0.917	0.702
	GM3	0.881	0.913	0.701
	GM4	0.912	0.952	0.682
	GM5	0.939	0.995	0.616
	GM6	0.937	0.990	0.641
10	No Grouping	0.908	0.958	0.624
	GM1	0.935	0.976	0.700
	GM2	0.935	0.965	0.763
	GM3	0.962	0.999	0.756
	GM4	0.926	0.973	0.657
	GM5	0.922	0.978	0.605
	GM6	0.928	0.997	0.536
15	No Grouping	0.934	0.991	0.615
	GM1	0.927	0.994	0.548
	GM2	0.945	0.988	0.704
	GM3	0.930	0.998	0.549
	GM4	0.927	0.990	0.570
	GM5	0.937	0.994	0.619
	GM6	0.928	0.996	0.539
20	No Grouping	0.909	0.971	0.561
	GM1	0.919	0.994	0.493
	GM2	0.919	0.992	0.500
	GM3	0.926	0.996	0.532
	GM4	0.918	0.971	0.615
	GM5	0.920	0.984	0.556
	GM6	0.937	0.996	0.604
25	No Grouping	0.926	0.988	0.575
	GM1	0.913	0.971	0.580
	GM2	0.904	0.966	0.554
	GM3	0.909	0.974	0.540
	GM4	0.931	0.997	0.556
	GM5	0.934	0.997	0.575
	GM6	0.921	0.996	0.496

Table 5.7: Agreement, sensitivity, and specificity of the HL model. The instances are solved with CPLEX 11.0 using a one-hour time limit.

Num. of	Grouping	Validation Set			Training Set			MIP gap%
Variables	Method	Agreement	Sensitivity	Specificity	Agreement	Sensitivity	Specificity	CPLEX
5	No Grouping	0.667	1.000	0.595	0.796	0.796	0.795	78.7
	GM1	0.734	0.647	0.753	0.765	0.774	0.750	23.2
	GM2	0.875	0.765	0.899	0.782	0.770	0.803	39.1
	GM3	0.880	0.882	0.880	0.804	0.796	0.818	42.5
	GM4	0.818	1.000	0.778	0.829	0.857	0.780	58.2
	GM5	0.880	0.882	0.880	0.840	0.891	0.750	64.7
	GM6	0.880	0.882	0.880	0.823	0.848	0.780	53.2
10	No Grouping	0.688	1.000	0.620	0.829	0.930	0.652	95.8
	GM1	0.823	1.000	0.785	0.890	0.913	0.848	92.6
	GM2	0.823	1.000	0.785	0.876	0.883	0.864	92.1
	GM3	0.745	1.000	0.690	0.859	0.870	0.841	93.1
	GM4	0.823	1.000	0.785	0.892	0.930	0.826	95.1
	GM5	0.823	1.000	0.785	0.870	0.917	0.788	95.5
	GM6	0.823	1.000	0.785	0.892	0.930	0.826	93.8
15	No Grouping	0.792	0.953	0.757	0.877	0.950	0.752	96.5
	GM1	0.779	0.953	0.742	0.860	0.904	0.782	94.5
	GM2	0.760	1.000	0.709	0.881	0.932	0.791	96.8
	GM3	0.806	1.000	0.765	0.890	0.943	0.797	95.8
	GM4	0.806	1.000	0.765	0.892	0.946	0.797	96.2
	GM5	0.802	1.000	0.759	0.882	0.934	0.791	96.1
	GM6	0.823	1.000	0.785	0.899	0.939	0.830	94.0
20	No Grouping	0.673	0.918	0.620	0.861	0.918	0.761	97.1
	GM1	0.775	0.953	0.737	0.861	0.913	0.770	94.9
	GM2	0.763	1.000	0.711	0.873	0.941	0.755	95.4
	GM3	0.806	1.000	0.765	0.893	0.953	0.788	95.0
	GM4	0.785	1.000	0.739	0.883	0.946	0.773	96.4
	GM5	0.846	0.953	0.823	0.893	0.953	0.788	94.8
	GM6	0.829	0.835	0.828	0.862	0.925	0.752	93.5
25	No Grouping	0.804	0.953	0.772	0.867	0.946	0.730	97.3
	GM1	0.758	1.000	0.706	0.863	0.906	0.788	93.7
	GM2	0.781	0.976	0.739	0.845	0.922	0.712	94.6
	GM3	0.794	1.000	0.749	0.877	0.946	0.758	94.9
	GM4	0.802	1.000	0.759	0.878	0.937	0.776	97.0
	GM5	0.806	1.000	0.765	0.896	0.965	0.776	95.4
	GM6	0.808	0.953	0.777	0.888	0.934	0.809	95.6

Table 5.8: Agreement, sensitivity, and specificity of the HL model on the data set of Smith et al. [171].

Num. of Variables	Grouping Method	Smith Set		
		Agreement	Sensitivity	Specificity
5	No Grouping	0.894	0.937	0.651
	GM1	0.890	0.951	0.546
	GM2	0.880	0.936	0.566
	GM3	0.908	0.952	0.654
	GM4	0.936	0.987	0.652
	GM5	0.919	0.979	0.585
	GM6	0.930	0.987	0.607
10	No Grouping	0.903	0.981	0.463
	GM1	0.927	0.998	0.524
	GM2	0.927	0.994	0.544
	GM3	0.916	0.980	0.554
	GM4	0.935	0.997	0.583
	GM5	0.933	0.996	0.572
	GM6	0.919	0.997	0.478
15	No Grouping	0.913	0.982	0.527
	GM1	0.911	0.985	0.496
	GM2	0.921	0.997	0.494
	GM3	0.928	0.996	0.543
	GM4	0.931	0.997	0.555
	GM5	0.928	0.995	0.552
	GM6	0.921	0.998	0.485
20	No Grouping	0.920	0.983	0.567
	GM1	0.913	0.988	0.485
	GM2	0.924	0.996	0.513
	GM3	0.924	0.997	0.507
	GM4	0.927	0.997	0.531
	GM5	0.920	0.996	0.493
	GM6	0.911	0.995	0.438
25	No Grouping	0.917	0.975	0.592
	GM1	0.909	0.986	0.476
	GM2	0.916	0.987	0.512
	GM3	0.925	0.997	0.513
	GM4	0.921	0.997	0.492
	GM5	0.921	0.998	0.482
	GM6	0.914	0.989	0.491

6.0 SUMMARY AND FUTURE RESEARCH

6.1 SUMMARY

This dissertation focuses on the annual influenza vaccine strain selection problem. It introduces multi-stage stochastic mixed-integer programming and bilevel stochastic mixed-integer programming models to analyze various aspects of the influenza vaccine preparation process. The primary goal is to increase societal vaccination benefit through integrating the composition and timing decisions of the flu shot design in a stochastic and dynamic environment.

In Chapter 3, we take the view of the Committee, and optimize strain selection decisions based on a production plan that is exogenously designed by the manufacturers. We exploit several structural properties, and propose a tailored branch-and-price algorithm. In addition, we incorporate risk-sensitivity through mean-risk models. We address multiple policy questions after carefully calibrating our models using publicly available data.

Our computational experiments provide valuable insights for pressing policy issues. In particular, we show that more frequent Committee meetings can provide up to 10% gains in the annual societal benefit of the flu shot. Incorporating more than three strains in the flu shot can increase the annual societal benefit by more than $80 million, particularly under more severe flu seasons. Finally, enhanced manufacturing techniques have substantial benefits, so that it may be in society's interest to subsidize research into such manufacturing techniques.

In Chapter 4, we extend the multi-stage stochastic mixed-integer programming model of Chapter 3 to consider the hierarchical relationship between the Committee and the influenza vaccine manufacturers. We formulate a two-level model for the annual influenza strain selection problem that incorporates both the society's and the vaccine manufacturers'

perspectives. We analyze the effects of yield uncertainty, price, and production cost on manufacturers' production decisions. We also analyze the properties of the consumer surplus using a hyperbolic demand function.

We find that if there is a single strain in the flu shot, then the expected industry output under yield uncertainty is always lower than that in the deterministic case. However, this is not true when there are multiple flu shot strains. Based on price of the vaccine, the expected industry output under yield uncertainty can exceed that in the deterministic case. The high flu shot price provides the manufacturer sufficient incentive to increase the production levels of strains in hope of a high level of realized yield.

Influenza viruses frequently mutate. If the mutant strains have significantly different antigenic properties, then the current influenza vaccine does not provide immunity protection and its composition has to be updated. Currently, hemagglutinin inhibition assays are used to identify antigenic variants of influenza A strains. However, this serology assay is labor-intensive and time-consuming. As an alternative, representing the difference between the antigenic properties of the influenza strains using a distance (i.e. antigenic distance) can be a rapid indicator of the likelihood that the current vaccine will protect against a recently emerged strain, and also facilitates the study of the virus' evolution.

In Chapter 5, we formulate cross validation as a bilevel program. Our model identifies amino acid substitutions that are positively associated with antigenic distances to predict antigenic variants of the influenza A/H3N2 virus. This approach allows for optimizing the cross validation outcome. Moreover, it offers modeling flexibility when considering multiple statistical learning goals including the average number of points misclassified, the maximum distance of each misclassified point from the classifier margin of its fold. We calibrate our model using serological data and compare our results with the current literature. Our models can identify potential immunodominant amino acid substitutions for predicting antigenic variants, so that could be readily integrated to the global influenza surveillance system.

6.2 LIMITATIONS AND FUTURE RESEARCH

The data sources that we mostly use include the Committee meeting materials [69], the CDC flu surveillance and vaccination reports [34], and also Morbidity and Mortality Weekly Reports [38]. However, published data on strain production are unavailable, so our numerical results should be interpreted cautiously.

Our models do not consider disease transmission, which is an important factor when assessing the potential benefits of vaccination. For example, by vaccinating an individual who would otherwise get infected, the infection cost of other people who would get the disease from that particular individual is also prevented. To this end, compartmental disease transmission models are more appropriate [118]. Our emphasis in this dissertation is more on addressing the optimal composition of the annual flu shot along with the timing of its manufacturing.

The effectiveness of the flu vaccine is different for every person. Moreover, the immune response to flu shot can extend more than one season [191]. One of the key factors determining how well the vaccine will protect is the individual's vaccination history [170]. On one hand, immunity gained from past flu vaccinations may cross–react with the new vaccine strains and reduce their immunogenicity. On the other hand, if the vaccine strain used in the past is antigenically similar to the currently circulating strains, then antibodies produced by the immune response to the past vaccine cross-react with the those strains and help clear the epidemic virus. Therefore, the efficacy of annually repeated vaccination remains unclear [19, 93, 102]. To capture the effect of repeated vaccination we could consider vaccination history as a parameter. However, taking into account each individual's vaccination history would lead to a policy which is impossible to implement. Instead, we can consider administering two types of vaccine with different contents. One is for people who were vaccinated last year, while the other is for people who were not vaccinated. This might lead to a notable increase in vaccine effectiveness at the expense of extra difficulties in the strain selection process.

We have assumed that uncertainty about future events is not affected or revealed by the current optimization decisions. Fewer people are expected to get infected by a flu strain

that is well matched by vaccine strains. Therefore strain selection decisions may affect the prevalence of flu types during the upcoming flu season. We can address this fact by formulating a stochastic programming model with endogenous uncertainty, where the underlying stochastic process depends on the strain selection decisions. Similar stochastic programming models with endogenous uncertainty are also considered in many stochastic programming models involving project design decisions over a time horizon [76, 77].

In Chapter 4, the Manufacturers' Problem (4.10) optimizes the expected profit. However, the manufacturers might be risk–averse when a making production decisions because of the yield uncertainty. We could address this issue using risk measures as we do for the strain selection problem in Section 3.2. For a review of risk measures used in stochastic programming, see [145, 164].

In Chapter 5 we use six different grouping methods for the amino acid sequence alignment. We can formulate a model that also develops a grouping method when optimizing the cross-validation outcome. In this model, there would be a binary variable for each amino acid pair, indicating whether those two amino acids should be classified as the same or not. Clearly, the number of such binary variables is exponential in the number of the amino acids. Therefore, we should analyze the model structure to develop efficient solution techniques.

BIBLIOGRAPHY

[1] S. Ahmed. Mean-risk objectives in stochastic programming. Technical report, Georgia Institute of Technology, 2004. Available from http://www2.isye.gatech.edu/∼sahmed/publications.html. Accessed September 16, 2010.

[2] S. Ahmed. Convexity and decomposition of mean-risk stochastic programs. *Mathematical Programming*, 106(3):433–446, 2006.

[3] S. Ahmed, M. Tawarmalani, and N. V. Sahinidis. A finite branch-and-bound algorithm for two-stage stochastic integer programs. *Mathematical Programming*, 100(2):355–377, 2004.

[4] O. Alagoz, L. M. Maillart, A. J. Schaefer, and M. S. Roberts. The optimal timing of living-donor liver transplantation. *Management Science*, 50(10):1420–1430, 2004.

[5] O. Alagoz, L. M. Maillart, A. J. Schaefer, and M. S. Roberts. Choosing among living-donor and cadaveric livers. *Management Science*, 53(11):1702–1715, 2007.

[6] O. Alagoz, L. M. Maillart, A. J. Schaefer, and M. S. Roberts. Determining the acceptance of cadaveric livers using an implicit model of the waiting list. *Operations Research*, 55(1):24–36, 2007.

[7] A. Alonso-Ayuso, L. F. Escudero, A. Garín, M. T. Ortuño, and G. Pérez. An approach for strategic supply chain planning under uncertainty based on stochastic $0-1$ programming. *Journal of Global Optimization*, 26(1):97–124, 2003.

[8] A. Alonso-Ayuso, L. F. Escudero, and M. T. Ortuño. BFC, a branch-and-fix coordination algorithmic framework for solving some types of stochastic pure and mixed $0-1$ programs. *European Journal of Operations Research*, 151(3):503–519, 2003.

[9] R. M. Anderson and R. M. May. *Infectious diseases of humans: dynamics and control*. Oxford University Press, New York, 1991.

[10] V. Andreasen and A. Sasaki. Shaping the phylogenetic tree of influenza by cross-immunity. *Theoretical Population Biology*, 70(2):164–173, 2006.

[11] C. Audet, P. Hansen, B. Jaumard, and G. Savard. Links between linear bilevel and mixed 0-1 programming problems. *Journal of Optimization Theory and Applications*, 93(2):273–300, 1997.

[12] J. F. Bard. *Practical bilevel optimization: algorithms and applications. Nonconvex optimization and its applications.* Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998.

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

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

[15] C. Barnhart, E. L. Johnson, G. L. Nemhauser, M. W. P. Savelsbergh, and P. H. Vance. Branch-and-price: Column generation for solving huge integer programs. *Operations Research*, 46(3):316–329, 1998.

[16] H. Behncke. Optimal control of deterministic epidemics. *Optimal Control Applications and Methods*, 21(6):269–285, 2000.

[17] O. Ben-Ayed. Bilevel linear programming. *Computers and Operations Research*, 20(5):485–501, 1993.

[18] R. S. Bernstein, D. C. Sokal, S. T. Seitz, B. Auvert, J. Stover, and W. Naamara. Simulating the control of a heterosexual HIV epidemic in a severely affected East African city. *Interfaces*, 28(3):101–126, 1998.

[19] W. E. Beyer, I. A. de Bruijn, A. M. Palache, R. G. Westendorp, and A. D. Osterhaus. Protection against influenza after annually repeated vaccination: A meta-analysis of serologic and field studies. *Archives of Internal Medicine*, 159(2):182–188, 1999.

[20] M. J. Brennan and T. M. Carroll. *Quantitative Economics and Econometrics*. South-Western Publishing Company, Cincinnati, USA, 1987.

[21] D. L. Brito, E. Sheshinski, and M. D. Intriligator. Externalities and compulsory vaccinations. *Journal of Public Economics*, 45(1):69–90, 1991.

[22] L. Brotcorne, M. Labbe, P. Marcotte, and G. Savard. A bilevel model for toll optimization on a multicommodity transportation network. *Transportation Science*, 35(4):1–14, 2001.

[23] R. M. Bush, C. A. Bender, K. Subbarao, N. J. Cox, and W. M. Fitch. Predicting the evolution of human influenza A. *Science*, 286(5446):1921–1925, 1999.

[24] R. M. Bush, W. M. Fitch, C. A. Bender, and N. J. Cox. Positive selection on the H3 hemagglutinin gene of human influenza virus A. *Molecular Biology and Evolution*, 16(11):1457–1465, 1999.

[25] C. C. Carøe and R. Schultz. Dual decomposition in stochastic integer programming. *Operations Research Letters*, 24(1):37–45, 1999.

[26] C. C. Carøe and J. Tind. L-shaped decomposition of two-stage stochastic programs with integer recourse. *Mathematical Programming*, 83(1-3):451–464, 1998.

[27] F. Carrat and A. Flahault. Influenza vaccine: the challenge of antigenic drift. *Vaccine*, 25(39-40):6852–6862, 2007.

[28] CDC. Delayed influenza vaccine availability for 2001-02 season and supplemental recommendations of the Advisory Committee on Immunization Practices. *Morbidity Mortality Weekly Reports*, 50:582–585, 2001.

[29] CDC. Interim influenza vaccination recommendations, 2004-2005 influenza season. *Morbidity Mortality Weekly Reports*, 53:923–924, 2004.

[30] CDC. Estimates of influenza vaccination target population sizes in 2006 and recent vaccine uptake levels, 2006. Available from http://www.cdc.gov/flu/professionals/vaccination/pdf/targetpopchart.pdf. Accessed September 16, 2010.

[31] CDC. The 2007–2008 flu season, 2008. Available from http://www.cdc.gov/flu/about/qa/season.htm. Accessed January 14, 2008.

[32] CDC. Influenza vaccine price list, 2008. Available from http://www.cdc.gov/vaccines/programs/vfc/downloads/archived-pricelists/2008/102008.htm. Accessed June 17, 2010.

[33] CDC. Interim within-season estimate of the effectiveness of trivalent inactivated influenza vaccine – Marshfield, Wisconsin, 2007–08 influenza season. *Morbidity & Mortality Weekly Reports*, 57(15):393–398, 2008.

[34] CDC. Flu activity & surveillance, 2010. Available from http://www.cdc.gov/flu/weekly/fluactivity.htm. Accessed September 16, 2010.

[35] CDC. Flu symptoms and severity, 2010. Available from http://www.cdc.gov/flu/about/disease/symptoms.htm. Accessed April 7, 2010.

[36] CDC. How well does the seasonal flu vaccine work? Seasonal influenza vaccine effectiveness, 2010. Available from http://www.cdc.gov/flu/about/qa/vaccineeffect.htm. Accessed April 7, 2010.

[37] CDC. Influenza viruses. 2010. Available from http://www.cdc.gov/flu/avian/gen-info/flu-viruses.htm. Accessed September 16, 2010.

[38] CDC. Morbidity & Mortality Weekly Reports, 2010. Available from http://www.cdc.gov/mmwr. Accessed September 16, 2010.

[39] CDC. Seasonal influenza: the disease, 2010. Available from http://www.cdc.gov/flu/about/disease/index.htm. Accessed September 16, 2010.

[40] CDC. Seasonal influenza vaccine supply for the U.S. 2009-10 influenza season, 2010. Available from http://www.cdc.gov/flu/about/qa/vaxsupply.htm. Accessed April 7, 2010.

[41] Centers for Medicare and Medicaid Services. National Health Expenditure Data. NHE Fact Sheet, 2009. Available from http://www.cms.gov/NationalHealthExpendData/25_NHE_Fact_Sheet.asp. Accessed June 15, 2011.

[42] J-P. Côté and G. Savard. A bilevel modelling approach to pricing and fare optimization in the airline industry. *Journal of Revenue and Pricing Management*, 2(1):23–26, 2003.

[43] J. Chhatwal, O. Alagoz, and E. S. Burnside. Optimal breast biopsy decision-making based on mammographic features and demographic factors. *Operations Research*, 58(6):1577–1591, 2010.

[44] S. E. Chick, H. Mamani, and D. Simchi-Levi. Supply chain coordination and influenza vaccination. *Operations Research*, 56(6):1493–1506, 2008.

[45] S.-H. Cho. The optimal composition of influenza vaccines subject to random production yields. *Manufacturing & Service Operations Management*, 12(2):256–277, 2010.

[46] B. Colson, P. Marcotte, and G. Savard. An overview of bilevel optimization. *Annals of Operations Research*, 153(1):235–256, 2007.

[47] N. J. Cox and K. Subbarao. Global epidemiology of influenza: past and present. *Annual Review of Medicine*, 51:407–421, 2000.

[48] ILOG CPLEX, 2010. Available from http://www.ilog.com/products/cplex/.

[49] G. B. Dantzig and G. Infanger. Multi-stage stochastic linear programs for portfolio optimization. *Annals of Operations Research*, 45(1):59–76, 1993.

[50] G. B. Dantzig and P. Wolfe. Decomposition principle for linear programs. *Operations Research*, 8(1):101–111, 1960.

[51] P. M. Danzon, N. S. Periera, and S. S. Tejwani. Vaccine supply: A cross national perspective. *Health Affairs*, 24(3):706–717, 2005.

[52] S. Dempe. Discrete bilevel optimization problems. Technical report, Institut fur Wirtschaftsinformatik, Universitat Leipzig, 2001.

[53] S. Dempe. *Foundations of bilevel programming.* Kluwer Academic Publishers, Dordrecht, The Netherlands, 2002.

[54] S. T. DeNegre and T. K. Ralphs. A branch-and-cut algorithm for bilevel integer programming. *Proceedings of the Eleventh INFORMS Computing Society Meeting*, pages 65–78, 2009.

[55] B. T. Denton, A. Miller, H. Balasubramanian, and T. Huschka. Optimal allocation of surgery blocks to operating rooms under uncertainty. *Operations Research*, 58(4):802–816, 2010.

[56] S. Deo and C. J. Corbett. Cournot competition under yield uncertainty: the case of the U.S. influenza vaccine market. *Manufacturing & Service Operations Management*, 11(4):563–576, 2009.

[57] G. Desaulniers, J. Desrosiers, and M. M. Solomon. Accelerating strategies in column generation methods for vehicle routing and crew scheduling problems. C. C. Ribeiro, P. Hansen, eds. *Essays and Surveys in Metaheuristics.* Kluwer, Boston, MA, pages 309–324, 2001.

[58] J. Desrosiers, F. Soumis, and M. Desrochers. Routing with time windows by column generation. *Networks*, 14(4):545–565, 1984.

[59] A. d'Onofrio, P. Manfredi, and E. Salinelli. Vaccinating behaviour, information, and the dynamics of SIR vaccine preventable diseases. *Theoretical Population Biology*, 71(3):301–317, 2007.

[60] W. D. D'Souza, R. R. Meyer, B. R. Thomadsen, and M. C. Ferris. An iterative sequential mixed-integer approach to automated prostate brachytherapy treatment plan optimization. *Physics in Medicine and Biology*, 46(2):297–322, 2001.

[61] O. du Merle, D. Villeneuve, J. Desrosiers, and P. Hansen. Stabilized column generation. *Discrete Mathematics*, 194(1-3):229–237, 1999.

[62] J. Dupačová. Portfolio optimization via stochastic programming: Methods of output analysis. *Mathematical Methods of Operations Research*, 50(2):245–270, 1999.

[63] G. D. Eppen, R. K. Martin, and L. Schrage. A scenario approach to capacity planning. *Operations Research*, 37(4):517–527, 1989.

[64] J. Espadaler, O. Romero-Isart, R. M. Jackson, and B. Oliva. Prediction of protein-protein interactions using distant conservation of sequence patterns and structure relationships. *Bioinformatics*, 21(16):3360–3368, 2005.

[65] S. Esposito, P. Marchisio, R. Droghetti, L. Lambertini, N. Faelli, S. Bosis, S. Tosi, E. Begliatti, and N. Principi. Influenza vaccination coverage among children with high-risk medical conditions. *Vaccine*, 24(24):5251–5255, 2006.

[66] N. D. Evans, L. J. White, M. J. Chapman, K. R. Godfrey, and M. J. Chappell. The structural identifiability of the susceptible–infected–recovered model with seasonal forcing. *Mathematical Biosciences*, 194(2):175–197, 2005.

[67] FDA. Vaccines and Related Biological Products Advisory Committee. Transcript for February 21, 2008, 2008. Available from http://www.fda.gov/ohrms/dockets/ac/cber08.html#VaccinesandRelatedBiological. Accessed April 10, 2010.

[68] FDA. Vaccines and Related Biological Products Advisory Committee. Transcript for February 18, 2009, 2009. Available from http://www.fda.gov/advisorycommittees/committeesmeetingmaterials. Accessed June 17, 2010.

[69] FDA. Vaccines and Related Biological Products Advisory Committee, 2010. Available from http://www.fda.gov/cber/advisory/vrbp/vrbpmain.htm. Accessed April 7, 2010.

[70] A. E. Fiore, D. K. Shay, K. Broder, J. K. Iskander, T. M. Uyeki, G. Mootrey, J. S. Bresee, and N. J. Cox. Prevention and control of influenza: Recommendations of the advisory committee on immunization practices. *Morbidity and Mortality Weekly Reports*, 57(1):1–60, 2008.

[71] L. R. Ford and D. R. Fulkerson. A suggested computation for maximal multicommodity network flows. *Management Science*, 5(1):97–101, 1958.

[72] J. Forrest, D. de la Nuez, and R. Lougee-Heimer. CLP user guide, 2004. Available from http://www.coin-or.org/Clp/userguide. Accessed July 2, 2010.

[73] General Accounting Office. Flu Vaccine: Supply Problems Heighten Need to Ensure Access for High–Risk People, 2001. Report to Congressional Requesters, General Accounting Office, Washington, DC. http://www.gao.gov/new.items/d01624.pdf.

[74] C. Gerdil. The annual production cycle for influenza vaccine. *Vaccine*, 21(16):1776–1779, 2003.

[75] GNU Scientific Library, 2011. Available from http://www.gnu.org/software/gsl/. Accessed June 15, 2011.

[76] V. Goel and I. E. Grossmann. A stochastic programming approach to planning of offshore gas field developments under uncertainty in reserves. *Computers and Chemical Engineering*, 28(8):1409–1429, 2004.

[77] V. Goel and I. E. Grossmann. A class of stochastic programs with decision dependent uncertainty. *Mathematical Programming*, 108(2-3):355–394, 2006.

[78] Y. Guan. Pairing inequalities and stochastic lot-sizing problems: a study in integer programming, 2005. Ph.D Dissertation, Georgia Institute of Technology.

[79] Y. Guan, S. Ahmed, and G. L. Nemhauser. Cutting planes for multi–stage stochastic integer programs. *Operations Research*, 57(2):287–298, 2009.

[80] R. W. Hamming. *Numerical Methods for Scientists and Engineers*. McGraw-Hill, New York, 1962.

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

[82] P. R. Harper and A. K. Shahani. Modelling for the planning and management of bed capacities in hospitals. *Journal of Operational Research Society*, 53(1):11–18, 2002.

[83] D.W. Hearn and M.V. Ramana. Solving congestion toll pricing models. In P. Marcotte, editor, *Equilibrium and advanced transportation modelling*, pages 109–124. Kluwer Academic Publishers, Boston, Massachusetts, 1998.

[84] T. Heinze and R. Schultz. A branch-and-bound method for multistage stochastic integer programs with risk objectives. *Optimization*, 57(2):277–293, 2008.

[85] H. Held and D. Woodruff. Heuristics for multi-stage interdiction of stochastic networks. *Journal of Heuristics*, 11(5-6):483–500, 2005.

[86] J. F. Hellinger. Expected utility theory and risky choices with health outcomes. *Medical Care*, 27(3):273–279, 1989.

[87] HHS. Contract to secure future egg supply for flu vaccines., 2004. Available from http://www.hhs.gov/news/press/2004pres/20041109a.html. Accessed April 9, 2010.

[88] HHS. Pandemic flu, 2008. Available from http://www.pandemicflu.gov. Accessed January 22, 2008.

[89] HHS. Vaccine production in cells. 2010. Available from http://www.flu.gov/professional/federal/vproductioncells.html. Accessed April 9, 2010.

[90] HIDA. 2008-2009 influenza vaccine production & distribution market brief, 2009. Available from http://www.flusupplynews.com/documents/09_FluBrief.pdf. Accessed June 17, 2010.

[91] J. C. Hornberger and J.-H. Ahn. Deciding eligibility for transplantation when a donor kidney becomes available. *Medical Decision Making*, 17(2):160–170, 1997.

[92] S. L. Hoshi, M. Kondo, Y. Honda, and I. Okubo. Cost-effectiveness analysis of influenza vaccination for people aged 65 and over in Japan. *Vaccine*, 25(35):6511–6521, 2007.

[93] T. W. Hoskins, J. R. Davies, A. J. Smith, C. L. Miller, and A. Allchin. Assessment of inactivated influenza-A vaccine after three outbreaks of influenza A at Christ's Hospital. *Lancet*, 313(1806):33–35, 1979.

[94] D. H. Howard. Why do transplant surgeons turn down organs? a model of theacceptreject decision. *Journal of Health Economics*, 21(6):957–969, 2002.

[95] K. Huang and S. Ahmed. The value of multi-stage stochastic programming in capacity planning under uncertainty. *Operations Research*, 57(4):893–904, 2009.

[96] S. A. Levin J. B. Plotkin, J. Dushoff. Hemagglutinin sequence clusters and the antigenic evolution of influenza A virus. *Proceedings of the National Academy of Sciences of the United States of America*, 99(9):6263–6268, 2002.

[97] H. Jin, H. Zhou, H. Liu, W. Chan, L. Adhikary, K. Mahmood, M. S. Lee, and G. Kemble. Two residues in the hemagglutinin of A/Fujian/411/02-like influenza viruses are responsible for antigenic drift from A/Panama/2007/99. *Virology*, 336(1):113–119, 2005.

[98] J. J. Júdice and A. M. Faustino. A sequential LCP method for bilevel linear programming. *Annals of Operations Research*, 34:89–106, 1992.

[99] E. P. C. Kao and M. Queyranne. Budgeting costs of nursing in a hospital. *Management Science*, 31(5):608–621, 1985.

[100] B. Y. Kara and V. Verter. Designing a road network for hazardous materials transportation. *Transportation Science*, 38(2):188–196, 2004.

[101] J. M. Katz and R. G. Webster. Efficacy of inactivated influenza A virus (H3N2) vaccines grown in mammalian cells or embryonated eggs. *Journal of Infectious Diseases*, 160(2):191–198, 1989.

[102] W. A. Keitel, T. R. Cate, R. B. Couch, L. L. Huggins, and K. R. Hess. Efficacy of repeated annual immunization with inactivated influenza virus vaccines over a 5 year period. *Vaccine*, 15(10):1114–1122, 1997.

[103] G. Kemble and H. Greenberg. Novel generations of influenza vaccines. *Vaccine*, 21(16):1789–1795, 2003.

[104] C. S. Khor, A. Elkamel, and P. L. Douglas. Stochastic refinery planning with risk management. *Petroleum Science and Technology*, 26(14):1726–1740, 2008.

[105] E. D. Kilbourne, B. E. Johansson, and B. Grajower. Independent and disparate evolution in nature of influenza A virus hemagglutinin and neuraminidase glycoproteins. *Proceedings of National Academy of Sciences of the United States of America*, 87:786–790, 1990.

[106] KNITRO, 2010. Ziena Optimization. Available from http://www.ziena.com/knitro.htm.

[107] S. Kodihalli, D. M. Justewicz, L. V. Gubareva, and R. G. Webster. Selection of a single amino acid substitution in the hemagglutinin molecule by chicken eggs can render influenza A virus (H3) candidate vaccine ineffective. *Journal of Virology*, 69(8):4888–4897, 1995.

[108] N. Kong, A. J. Schaefer, B. K. Hunsaker, and M. S. Roberts. Maximizing the efficiency of the U. S. liver allocation system through region design. *Management Science*, 56(12):2111–2122, 2010.

[109] L. J. Kornish and R. L. Keeney. Repeated commit-or-defer decisions with a deadline: the influenza vaccine composition. *Operations Research*, 56(3):527–541, 2008.

[110] P. Kotler. *Marketing decision making: a model building approach*. Holt, Rinehart and Winston, New York, 1971.

[111] J. Kruskal. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometika*, 29(1):1–27, 1964.

[112] J. Kruskal. Nonmetric multidimensional scaling. *Psychometika*, 29(2):115–129, 1964.

[113] S. Kuhn and R. Schultz. Risk neutral and risk averse power optimization in electricity networks with dispersed generation. *Mathematical Methods of Operations Research*, 69(2):353–367, 2009.

[114] G. Kunapuli, K. P. Bennett, J. Hu, and J-S. Pang. Classification model selection via bilevel programming. *Optimization Methods and Software*, 23(4):475–489, 2008.

[115] H. C. Kung, D. L. Hoyert, J. Xu, and S. L. Murphy. Deaths: final data for 2005. *National Vital Statistical Reports*, 56(10):1–120, 2008.

[116] L. Lambertini. Oligopoly with hyperbolic demand: A differential game approach. *Journal of Optimization Theory and Applications*, 145(1):108–119, 2010.

[117] A. Lapedes and R. Farber. The geometry of shape space: application to influenza. *Journal of Theoretical Biology*, 212(1):57–69, 2001.

[118] R. C. Larson. Simple models of influenza progression within a heterogeneous population. *Operations Research*, 55(3):399–412, 2007.

[119] D. P. Laurie. Calculation of Gauss-Kronrod quadrature rules. *Mathematics of Computation*, 66(219):1133–1145, 1997.

[120] E. K. Lee, T. Fox, and I. Crocker. Integer programming applied to intensity-modulated radiation therapy treatment planning-application to permanent prostate implants. *Annals of Operations Research*, 119:165–181, 2003.

[121] E. K. Lee and M. Zaider. Mixed integer programming approaches to treatment planning for brachytherapy. *Annals of Operations Research*, 119:147–163, 2003.

[122] H-Y. Lee, D. J. Topham, S-Y. Park, J. Hollenbaugh, J. Treanor, T. R. Mosmann, X. Jin, B. M. Ward, H. Miao, J. Holden-Wiltse, A. S. Perelson, M. Zand, and H. Wu. Simulation and prediction of the adaptive immune response to influenza A virus infection. *Journal of Virology*, 83(14):7151–7165, 2009.

[123] M-S. Lee and J-S. Chen. Predicting antigenic variants of influenza A/H3N2 viruses. *Emerging Infectious Diseases*, 10(8):1385–1390, 2004.

[124] M-S. Lee, M-C. Chen, Y-C. Liao, and C. A. Hsiung. Identifying potential immunodominant positions and predicting antigenic variants of influenza A/H3N2 viruses. *Vaccine*, 25(48):8133–8139, 2007.

[125] C. Lefévre. Optimal control of a birth and death epidemic process. *Operations Research*, 29(5):971982, 1981.

[126] Y-C. Liao, M-S. Lee, C-Y. Ko, and C. A. Hsiung. Bioinformatics models for predicting antigenic variants of influenza A/H3N2 virus. *Bioinformatics*, 24(4):505–512, 2008.

[127] P. Loridan and J. Morgan. Weak via strong Stackelberg problem: new results. *Journal of Global Optimization*, 8(3):263–287, 1996.

[128] M. E. Lübbecke and J. Desrosiers. Selected topics in column generation. *Operations Research*, 53(6):1007–1023, 2005.

[129] G. Lulli and S. Sen. A branch-and-price algorithm for multistage stochastic integer programming with application to stochastic batch-sizing problems. *Management Science*, 50(6):786–796, 2004.

[130] A. Madansky. Inequalities for stochastic linear programming problems. *Management Science*, 6(2):197–204, 1960.

[131] L. M. Maillart, J. S. Ivy, S. Ransom, and K. Diehl. Assessing dynamic breast cancer screening policies. *Operations Research*, 56(5):1411–1427, 2008.

[132] P. Marcotte and G. Savard. *Bilevel programming: a combinatorial perspective. In: D. Avis, A. Hertz, and O. Marcotte (Eds.). Graph theory and combinatorial optimization.* Boston: Kluwer Academic Publishers, 2005.

[133] A. Märkert and R. Schultz. On deviation measures in stochastic integer programming. *Operations Research Letters*, 33(5):441–449, 2005.

[134] H. M. Markowitz. *Portfolio selection: efficient diversification of investments.* John Wiley & Sons, 1959.

[135] M. I. Meltzer, K. M. Neuzil, M. R. Griffin, and K. Fukuda. An economic analysis of annual influenza vaccination of children. *Vaccine*, 23(8):1004–1014, 2005.

[136] R. R. Meyer, W. D. D'Souza, M. C. Ferris, and B. R. Thomadsen. MIP models and B&B strategies in brachytherapy treatment optimization. *Journal of Global Optimization*, 25(1):23–42, 2003.

[137] A. Migdalas. Bilevel programming in traffic planning: models, methods and challenge. *Journal of Global Optimization*, 7(4):381–405, 1995.

[138] A. Migdalas, P. M. Pardalos, and P. Värbrand. *Multilevel optimization: algorithms and applications.* Norwell: Kluwer Academic Publishers, 1998.

[139] H. E. Miller, W. P. Pierskalla, and G. J. Rath. Nurse scheduling using mathematical programming. *Operations Research*, 24(5):857–870, 1976.

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

[141] E. T. Muñoz and M. W. Deem. Epitope analysis for influenza vaccine design. *Vaccine*, 23(9):1144–1148, 2005.

[142] L. Murphy, R. J. Kaye, and F. F. Wu. Distributed spot pricing in radial distribution systems. *IEEE Transactions on Power Systems*, 9(1):311–317, 1994.

[143] R. W. Newman, R. Jennings, D. L. Major, J. S. Robertson, R. Jenkins, C. W. Potter, I. Burnett, L. Jewes, M. Anders, D. Jackson, and J. S. Oxford. Immune response of human volunteers and animals to vaccination with egg-grown influenza A (H1N1) virus is influenced by three amino acid substitutions in the haemagglutinin molecule. *Vaccine*, 11(4):400–406, 1993.

[144] K. L. Nichol, K. P. Mallon, and P. M. Mendelman. Cost benefit of influenza vaccination in healthy, working adults: an economic analysis based on the results of a clinical trial of trivalent live attenuated influenza virus vaccine. *Vaccine*, 21(17-18):2207–2217, 2003.

[145] W. Ogryczak and A. Ruszczyński. Dual stochastic dominance and related mean–risk models. *SIAM Journal of Optimization*, 13(1):60–78, 2002.

[146] R. Omori, B. Adams, and A. Sasaki. Coexistence conditions for strains of influenza with immune cross-reaction. *Journal of Theoretical Biology*, 262(1):48–57, 2010.

[147] O. Y. Özaltın, O. A. Prokopyev, and A. J. Schaefer. The bilevel knapsack problem with stochastic right-hand sides. *Operations Research Letters*, 38(4):328–333, 2010.

[148] O. Y. Özaltın, O. A. Prokopyev, A. J. Schaefer, and M. S. Roberts. Optimizing the societal benefits of the annual influenza vaccine: a stochastic programming approach, 2010. Accepted for publication in *Operations Research*.

[149] J. Patrick, M. L. Puterman, and M. Queyranne. Dynamic multi-priority patient scheduling. *Operations Research*, 56(6):1507–1525, 2008.

[150] A. S. Perelson and G. F. Oster. Theoretical studies of clonal selection: minimal antibody repertoire size and reliability of self-non-self discrimination. *Journal of Theoretical Biology*, 81(4):645–670, 1979.

[151] W. P. Pierskalla and D. J. Brailer. *Applications of Operations Research in Healthcare Delivery.* In: Handbooks in Operations Research and Management Science, 6: Operations Research and the Public Sector, S. M. Pollock, M. H. Rothkopf, A. Barnett (Eds.), Elsevier, Amsterdam, The Netherlands, 1994.

[152] M. J. D. Powell. *A Hybrid Method for Nonlinear Equations.* In: Numerical Methods for Nonlinear Algebraic Equations, P. Rabinowitz (Eds.), Gordon and Breach, New York, 1970.

[153] F. Preciado-Walters, R. Rardin, M. Langer, and V. Thai. A coupled column generation, mixed integer approach to optimal planning of intensity modulated radiation therapy for cancer. *Mathematical Programming (Series B)*, 101(2):319–338, 2004.

[154] J. R. Quinlan, 1993. C4.5: Programs for Machine Learning. San Mateo, CA: Morgan Kaufmann.

[155] T. K. Ralphs and L. Ladányi. COIN/BCP user's manual, 2001. Available from http://www.coin-or.org/Presentations/bcp-man.pdf. Accessed July 2, 2010.

[156] R. Rappuoli. Cell-culture-based vaccine production: technological options. *The Bridge*, 36(3):25–30, 2006.

[157] A. Richter, M. L. Brandeau, and D. K. Owens. An analysis of optimal resource allocation for HIV prevention among injection drug users and nonusers. *Medical Decision Making*, 19(2):167–179, 1999.

[158] E. Rizakow, J. Rosenhead, and K. Reddington. AIDSPLAN: A decision support model for planning the provision of HIV/AIDS related services. *Interfaces*, 21(3):117–129, 1991.

[159] H. E. Romeijn, R. K. Ahuja, J. F. Dempsey, and A. Kumar. A new linear programming approach to radiation therapy treatment planning problems. *Operations Research*, 54(2):201–216, 2006.

[160] W. Römisch and R. Schultz. *Multistage stochastic integer programs: an introduction.* In: Online Optimization of Large Scale Systems, M. Grötschel and S. O. Krumke and J. Rambau (Eds.), Springer, Berlin, 2001.

[161] W. Rudin. *Principles of mathematical analysis.* International series in pure and applied mathematics. McGraw-Hill, 1976.

[162] A. Ruszczyński and A. Shapiro. Optimization of convex risk functions. *Mathematics of Operations Research*, 31(3):433–452, 2006.

[163] B. Sandıkçı, L. M. Maillart, A. J. Schaefer, O. Alagoz, and M. S. Roberts. Estimating the patient's price of privacy in liver transplantation. *Operations Research*, 56(6):1393–1410, 2008.

[164] R. Schultz. Stochastic programming with integer variables. *Mathematical Programming*, 97(1–2):285–309, 2003.

[165] R. Schultz, L. Stougie, and M. H. van der Vlerk. Solving stochastic programs with integer recourse by enumeration: a framework using Gröbner basis reductions. *Mathematical Programming*, 83(1-3):229–252, 1998.

[166] R. Schultz and S. Tiedemann. Conditional value-at-risk in stochastic programs with mixed-integer recourse. *Mathematical Programming*, 105(2-3):365–386, 2006.

[167] E. C. Sewell and S. H. Jacobson. Using an integer programming model to determine the price of combination vaccines for childhood immunization. *Annals of Operations Research*, 119(1-4):261–284, 2003.

[168] D. M. Shepard, M. C. Ferris, G. H. Olivera, and T. R. Mackie. Optimizing the delivery of radiation therapy to cancer patients. *SIAM Review*, 41(4):721–744, 1999.

[169] K. J. Singh, A. B. Philpott, and R. K. Wood. Dantzig-Wolfe decomposition for solving multi-stage stochastic capacity-planning problems. *Operations Research*, 57(5):1271–1286, 2009.

[170] D. J. Smith, S. Forrest, D. H. Ackley, and A. S. Perelson. Variable efficacy of repeated annual influenza vaccination. *Proceedings of the National Academy of Sciences of the United States of America*, 96(24):14001–14006, 1999.

[171] D. J. Smith, A. S. Lapedes, J. C. de Jong, T. M. Bestebroer, G. F. Rimmelzwann, A. D. M. E. Osterhaus, and R. A. M. Fouchier. Mapping the antigenic and genetic evolution of influenza virus. *Science*, 305(5682):371–376, 2004.

[172] J. E. Stahl, N. Kong, S. M. Shechter, A. J. Schaefer, and M. S. Roberts. A methodological framework for optimally reorganizing liver transplant regions. *Medical Decision Making*, 25(1):35–46, 2005.

[173] R. B. Statnikov and J. B. Matusov. *Multi criteria optimization and engineering*. Chapman & Hall, New York, 1995.

[174] J. Stoer and R. Bulirsch. *Introduction to Numerical Analysis*. Springer, ISBN 978-0-387-95452-3 (3rd ed.), 2002.

[175] S. Takriti and S. Ahmed. On robust optimization of two-stage systems. *Mathematical Programming*, 99(1):109–126, 2004.

[176] S. Takriti and J. R. Birge. Lagrangian solution techniques and bounds for loosely coupled mixed-integer stochastic programs. *Operations Research*, 48(1):91–98, 2000.

[177] S. Takriti, J. R. Birge, and E. Long. A stochastic model for the unit commitment problem. *IEEE Transactions on Power Systems*, 11(3):1497–1508, 1996.

[178] M. W. Tanner, L. Sattenspiel, and L. Ntaimo. Finding optimal vaccination strategies under parameter uncertainty using stochastic programming. *Mathematical Biosciences*, 215(2):144–151, 2008.

[179] F. Vanderbeck. On Dantzig-Wolfe decomposition in integer programming and ways to perform branching in a branch-and-price algorithm. *Operations Research*, 48(1):111–128, 2000.

[180] F. Vanderbeck and L. A. Wolsey. An exact algorithm for IP column generation. *Operations Research Letters*, 19(4):151–159, 1996.

[181] T. Vu, S. Farish, M. Jenkins, and H. Kelly. A meta-analysis of effectiveness of influenza vaccine in persons aged 65 years and over living in the community. *Vaccine*, 20(13-14):1831–1836, 2002.

[182] U. P. Wen and Y. H. Yang. Algorithms for solving the mixed integer two-level linear programming problem. *Computers and Operations Research*, 17(2):133–142, 1990.

[183] D. Weycker, J. Edelsberg, M. E. Halloran, I. M. Longini, A. Nizam, V. Ciuryla, and G. Oster. Population-wide benefits of routine vaccination of children against influenza. *Vaccine*, 23(10):1284–1293, 2005.

[184] WHO. Recommendations for influenza vaccine composition., 2010. http://www.who.int/csr/disease/influenza/vaccinerecommendations1/en/. Accessed April 7, 2010.

[185] WHO. Seasonal influenza, 2010. Available from http://www.who.int/mediacentre/factsheets/fs211/en/index.html Accessed September 15, 2010.

[186] I. A. Wilson and N. J. Cox. Structural basis of immune recognition of influenza virus hemagglutinin. *Annual Review of Immunology*, 8:737–771, 1990.

[187] J. M. Wood, J. S. Oxford, U. Dunleavy, R. W. Newman, D. Major, and J. S. Robertson. Influenza A (H1N1) vaccine efficacy in animal models is influenced by two amino acid substitutions in the hemagglutinin molecule. *Virology*, 171(1):214–221, 1989.

[188] R. K. Wood. Deterministic network interdiction. *Mathematical and Computer Modelling*, 17(2):1–18, 1993.

[189] World Health Organization. The World Health Report, 2000. Available from http://www.who.int/whr/2000/en/whr00_en.pdf. Accessed June 15, 2011.

[190] World Health Organization. WHO Statistical Information System. World Health Statistics, 2011. Available from http://www.who.int/whosis/whostat/2011/en/index.html. Accessed June 15, 2011.

[191] J. T. Wu, L. M. Wein, and A. S. Perelson. Optimization of influenza vaccine selection. *Operations Research*, 53(3):456–476, 2005.

[192] J. M. Yusta, H. M. Khodr, and A. J. Urdaneta. Optimal pricing of default customers in electrical distribution systems: Effect behavior performance of demand response models. *Electrical Power Systems Research*, 77(5-6):548–558, 2007.

[193] F. Zhang and Z. Liv. Global stability of an SIR epidemic model with constant infectious period. *Applied Mathematics and Computation*, 199(1):285–291, 2008.