OPERATIONS RESEARCH
Vol. 56, No. 3, May–June 2008, pp. 527–541
ISSN 0030-364X | EISSN 1526-5463 | 08 | 5603 | 0527

informs

DOI 10.1287/opre.1070.0436

Repeated Commit-or-Defer Decisions with a Deadline: The Influenza Vaccine Composition

Laura J. Kornish
Leeds School of Business, University of Colorado, Boulder, Colorado 80309, kornish@colorado.edu

Ralph L. Keeney
Fuqua School of Business, Duke University, Durham, North Carolina 27708, keeney@duke.edu

Seasonal products have an effective inventory deadline, a time by which the inventory must be ready to distribute. The deadline creates an incentive to start early with production. However, opportunities to gather information that might change production decisions provide an incentive to defer the start of production. We study the resultant dynamic decision problem with alternatives that commit to one of several courses of action now and an alternative to defer the commitment to gather more information about the possible consequences of each alternative. The deadline increases the effective cost of gathering information because that cost includes the value sacrificed by reducing the time available to produce inventory. We frame our model using the annual influenza vaccine composition decision: deciding between strains of the virus to include, which must happen in the spring to allow time for vaccine production before the fall flu season begins. Our analysis describes the optimal decision strategies for this commit-or-defer decision. Many insights are drawn from this model that could contribute to more informed flu vaccine composition decisions. We comment on the relevance of this commit-or-defer decision model to a firm's production decisions for other seasonal products with an inventory deadline such as fashion goods.

Subject classifications: decision analysis: sequential; dynamic programming/optimal control: applications; government: regulations.
Area of review: Decision Analysis.
History: Received January 2006; revision received July 2006; accepted November 2006. Published online in *Articles in Advance* October 15, 2007.

1. Introduction

Many important decisions involve a choice among several alternatives, with the option to defer the choice to learn more. These "commit-or-defer" choices are ubiquitous in business, government, and individual decisions.

This paper examines an important class of commit-or-defer decisions involving a time deadline by which a definitive choice must be made and implemented. A classic example is the annual public sector choice of the influenza vaccine composition by the Food and Drug Administration (FDA). Distinguishing features of this problem include an inventory deadline, a product useful for one season only, a time-consuming production process, and the possibility of gathering some information about the appropriateness or potential value of the product configurations. Because of this deadline, the cost of information gathering includes not only its monetary cost, but also the value sacrificed by reducing the interval for production. Choosing early allows time to produce all of the inventory that is desired, but the product may not turn out to be the one most useful. Choosing later enhances the likelihood of producing the right product, but the inventory may be inadequate because of the reduced production time. The fundamental trade-off is to balance having the right inventory and having enough inventory.

We use this influenza vaccine composition decision to develop and illustrate a model to analyze repeated commit-or-defer decisions. This work has a decision analysis perspective: defining the decision criteria, using the dynamic nature of the problem in developing decision strategies, and incorporating the important uncertainties in the problem. The model lends insights into the dynamic decision problem, and we (1) derive the form of the optimal solution and present results about how the optimal actions change with changes in the deadline and the production rate, (2) use these results to make recommendations for the flu vaccine composition decision, and (3) describe the implications for the trade-offs between immediate action and delay for commit-or-defer decisions under broad assumptions.

This paper is organized as follows. Section 2 summarizes literature relevant to commit-or-defer decisions. Section 3 discusses the importance of the decision about the influenza vaccine composition and outlines the decision process used to make this annual decision in the United States. Section 4 presents our basic model of this decision, which is analyzed in §5. Section 6 presents a qualitative sensitivity analysis of the assumptions and parameters in the model to provide insights for the vaccine decision process. The relevance of the basic model to other commit-or-defer

decisions is discussed in §7. Section 8 contains concluding remarks.

2. Literature

Many classes of sequential decision problems under uncertainty with the opportunity to gather information between decisions have been studied. Work in sequential analysis (see e.g., Girshick 1946, Wald 1947, DeGroot 1970, and Chow et al. 1971) addresses problems in which information is gathered prior to each step where a decision about whether or not to proceed is made. Wald (1947) shows that under general conditions, this type of stopping problem is characterized by thresholds: when the information points convincingly in one direction, a "stopping" action is optimal, otherwise "continue gathering information" is optimal. DeGroot (1970) devotes a significant part (§4) of his classic Bayesian decision theory book to sequential decisions. Particularly relevant to our work are the sections of DeGroot (1970, pp. 277 ff.) and Chow et al. (1971, pp. 49 ff.) that discuss finite-horizon sequential analysis problems and the technique of backward induction. They show how the finite-horizon problems are useful for building successive approximations for infinite-horizon problems, i.e., those in which there is not an imposed upper bound on the number of samples. In the problem we study, the seasonal deadline provides a naturally occurring finite horizon.

Another well-studied sequential decision problem is the problem of search. Search is a broad term for problems in which one of the currently available alternatives can be selected or the decision can be deferred to locate new alternatives or new information. These problems are cast in a variety of ways, including search for the best price (e.g., Rothschild 1974) or the best job (e.g., Lippman and McCall 1976, Mortensen 1986). While the standard formulation of the problem is infinite horizon, Lippman and McCall (1976) analyze a finite-horizon search (pp. 166–171) as does Mortensen (1986, pp. 860–861).

Mortensen distinguishes the standard search problem, which is a "search for new offers" problem, from the "learning about the job" problem. While the "search for new offers" problem can be thought of as information gathering (as in information gathering about the "locations" of new job opportunities), it is not the same kind of information gathering about forecasts and distributions that we have in our model. The "learning about the job" problem is closer to our information gathering formulation: the job is treated as an "experience good" rather than an "inspection good." The worker learns about a job as he works at it, and if his assessment of future prospects falls below some threshold, then he will dip back into the job pool for another draw.

Models of technology choice are also cast as sequential decision problems. Jensen (1982) and McCardle (1985) analyze adoption decisions for a technology of unknown profitability. In Jensen's model, information is costlessly observed each period; in McCardle's model there is a fixed cost of observation per unit of time with the option to stop the observation process altogether. McCardle explicitly discusses the analysis of a single new technology, but implicitly analyzes a choice between something new and a fallback option, with known value. Lippman and McCardle (1991) combine "search" (efforts to identify new alternatives) and information gathering (learning about the distribution for a particular alternative). The three main differences compared to our paper are as follows. In their work, alternatives are considered one at a time instead of simultaneously, as in our work; there may be uncountably many alternatives as opposed to finitely many; and there is no deadline.

The repeated commit-or-defer decision we study in this paper shares some aspects of the technology adoption problem in the face of competing standards studied by Kornish (2006). In both models, deferring a decision about which alternative to adopt allows more time to obtain information about which of the competing contenders will dominate. One important difference in the present work is the deadline in the problem (the start of flu season). Our model applies to situations where the decision maker is observing trends about the "relevance" of different possible products, but must stop observing and commit production capacity to one product because of the inventory deadline.

The vaccine composition decision is closely related to other commit-or-defer decisions, such as those concerning fashions for a new season (see, e.g., Fisher et al. 1994 and Fisher and Raman 1996). Monitoring the prevalence of viral strains is similar to tracking the early-season popularity of different styles or the popularity of different telecommunications standards. The two problems differ somewhat, however, in the way the benefits of the produced units accrue, which affects the objective function of the analysis. With markets for goods, value is created by satisfying the demand of customers; the customers who benefit identify themselves by revealing their willingness to buy the product. With vaccines, the beneficiaries are statistical because they can never be identified. They are the people who do not contract the illness, but would have gotten the illness if they had not been vaccinated.

Similar in spirit to our analytical approach is the work of Wu et al. (2005). They develop a theory of antigenic distance and propose an optimization of vaccine selection for an individual based on the individual's and the population-level vaccination histories. In contrast, our approach focuses on a population-level recommendation based on tracking the spread of the candidate strains in the current season.

3. The Influenza Vaccine Composition Decision

The influenza vaccine decision used to develop and illustrate our analysis is a high-profile, important, classic commit-or-defer problem in the public sector affecting millions

of people in the United States each year. Over 80 million people were vaccinated for the 2003–2004 season, and annual influenza deaths average 36,000 (Thompson et al. 2004).

Treanor (2004) explains why the changing nature of the circulating flu strains makes the vaccine supply chain particularly vulnerable. The last century's pandemics (1918, 1957, and 1968) and close calls (e.g., 1977) bear out his assertion. Neustadt and Fineberg (1982) recount the federal government's vaccine decisions in the 1976–1977 season. They include an appendix of questions to guide vaccine policy decisions, indicating that a key question to ask is "What new information would cause you to change some or all of the recommendations you have made?" (p. 221). This contingent type of thinking is at the heart of our model.

The flu vaccine composition decision made headlines in the spring of 2003 because of the emergence of the new Fujian strain (CNN.com 2003). With the new strain, there was much discussion about how widespread it would become and whether or not there was time to develop and produce a vaccine before the start of the fall flu season (the inventory deadline in our formulation). During the 2004–2005 flu season, vaccine production was again in the news for two reasons. First, contamination at Chiron's Liverpool facility left the United States with half the anticipated vaccine supply (Whalen et al. 2004). Second, the specter of a new strain of avian flu has the World Health Organization (WHO) (CNN.com 2004) and the U.S. Center for Disease Control and Prevention (CDC.gov 2006) concerned about a pandemic.

In the United States and much of the world, the influenza vaccine is a trivalent vaccine, including one strain of each of three categories of virus (labeled A1, A2, and B). Influenza viruses are in a perpetual state of flux, with changes happening in large and small steps (antigenic "shifts" and "drifts," respectively). Due to these changes, early in each year (in recent years in February), the Vaccines and Related Biologic Products Advisory Committee (VRBPAC) recommends to the FDA which strains of each flu virus to include in the vaccine to be delivered starting in the fall. The extensive public record and transcripts of the official committee's deliberation about vaccine production (VRBPAC 2003, 2004, 2005) offers us an opportunity to scrutinize this decision-making process. In some years, the VRBPAC recommends a new strain for inclusion. Because of the annual time frame, if a new strain is included, the FDA does not require the customary clinical trials for the new vaccine: the license to produce the trivalent vaccine is applicable across the annual changes. Table 1 summarizes the strains that have been chosen in each of the categories in recent seasons. In three out of the last four seasons, there have been changes in at least one component of the vaccine.

When the committee meets to discuss the recommendation, there is an option to defer the recommendation to a later time when more information would be available. However, because the virus is grown in eggs and the manufacturing timeline includes many stages with safety and efficacy tests, vaccine production is time intensive. This forces the vaccine composition decision to be made early in the year and any deferral leaves less time for production before the start of flu season. Because the required production time is often a binding constraint on the problem, there is a trade-off between quantity (producing more) and quality (produce a more effective vaccine because you know more).

By deferring the decision, information can be gathered to reduce uncertainty about the coming flu season. There are several sources of uncertainty in this decision, including the anticipated prevalence of each strain of the virus, production issues for each strain, and effectiveness of vaccines against flu caused by the different strains. Reduced uncertainty results in a more-informed decision, which should lead to a more effective vaccine. In our analysis, we concentrate on the first type of uncertainty: the size of the population that would be stricken by each strain in the absence of the vaccine. The option to defer can be evaluated by analyzing the way in which potential information revelation would change the preferred action.

A report from the Institute of Medicine (IOM 2004) explains how the prominent role of the government as regulator (and for some vaccines, although not generally for flu, the purchaser) of vaccines has made the supply for all the recommended vaccines quite fragile. The closure of Chiron's plant in 2004 illustrates this fragility. Currently, there are only a handful of authorized influenza vaccine producers for the United States: Sanofi Pasteur, Chiron, and MedImmune, and the latest addition, GlaxoSmithKline. This paucity demonstrates that this market is not a terribly attractive one for the producers. The IOM report argues that legal and regulatory issues stifle innovation and dampen financial incentives in the vaccine markets.

4. Model

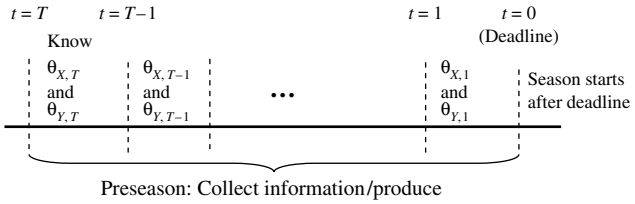
We develop a general model in this section that captures the repeated nature of the decision over time: either com-

Table 1. The strain selection for each of the three categories of virus—A1, A2, and B.

	2002–2003	2003–2004	2004–2005	2005–2006
A1	A/New Caledonia/20/99	A/New Caledonia/20/99	A/New Caledonia/20/99	A/New Caledonia/20/99
A2	A/Moscow/10/99	A/Moscow/10/99	**A/Fujian/411/2002**	**A/California/7/2003**
B	**B/Hong Kong/330/2001**	B/Hong Kong/330/2001	**B/Shanghai/361/2002**	B/Shanghai/361/2002

Source. FDA website (VRBPAC 2003, 2004, 2005). Changes from the previous year are indicated in boldface.

Figure 1. Timeline for repeated commit-or-defer decisions.



mit to a course of action now or defer this definitive choice to gather more information. We consider a discrete-time model, in which the decision is considered at the beginning of every period. The problem has a natural finite horizon because of the deadline. Figure 1 shows a timeline for the model, where t indicates the time periods remaining before the deadline at $t = 0$ and the first decision occurs at $t = T$, with the information available about each strain represented by $\theta_{\cdot,t}$. In each period, the decision maker has a choice among two definitive alternatives and deferring this decision to the next period. In the vaccine context, the definitive alternatives are to select one of two strains of the virus considered for one of the virus categories. For this context, it is useful to think of t measured in weeks and T occurring in February with the flu season beginning (i.e., $t = 0$) October 1.

There are two equivalent ways to state the objective function to measure the value of alternative choices for our model. One is to minimize the expected number of flu cases that occur in the coming flu season. The other is to maximize the number of flu cases that are prevented by using the vaccine produced. This objective implicitly assumes that all flu cases are equivalent, even when they may be caused by a different strain of a virus. It also does not separately account for the number of flu deaths prevented by the vaccine, the total cost of the program, or potential side effects from administering the vaccine. These issues are considered in the sensitivity analyses discussed in §6.

The first two decision periods in our model are represented in Figure 2, which also serves to illustrate our nota-

Figure 2. Decision tree for the initial two decision periods.

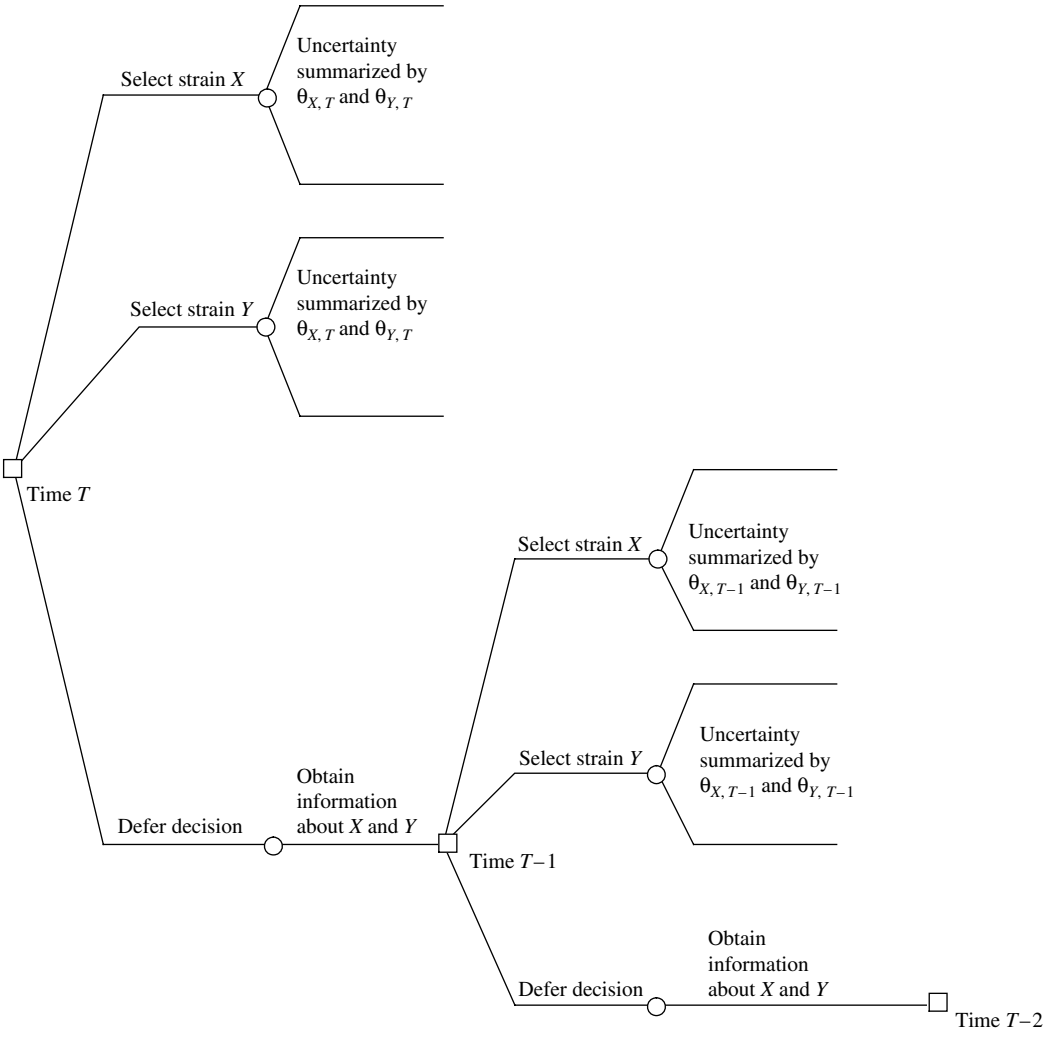


Table 2. Notation summary.

r	Production rate of vaccine.
t	Time periods remaining before deadline; $t = 0, 1, \ldots, T$.
z	Size of the U.S. population.
p	Percentage of population that seeks vaccination.
$\theta_{X,t}$	Summary of the information available at period t to estimate x_t.
$\theta_{Y,t}$	Summary of the information available at period t to estimate y_t.
x_t	Assessment of the mean number of cases of flu due to strain X that would occur during the season if no vaccine is used, determined at time t and based on $\theta_{X,t}$. x_t is the time-t forecast for strain X.
y_t	Assessment of the mean number of cases of flu due to strain Y that would occur during the season if no vaccine is used, determined at time t and based on $\theta_{Y,t}$. y_t is the time-t forecast for strain Y.
V_t	Value of optimal decisions from period t until deadline.

tion summarized in Table 2. At time T, we can select either strain X or strain Y of the virus or defer that decision to time $T-1$. With T periods remaining, there are uncertainties about the number of cases of the flu that would be caused by strain X and by strain Y in the coming season if there were no vaccine. The information about strains X and Y at time T is summarized by $\theta_{X,T}$ and $\theta_{Y,T}$, respectively, with time-T point-estimate forecasts denoted x_T and y_T for the expected flu cases in the coming season from each strain if no vaccine were administered. If a decision is deferred at time T, information is learned about the possible prevalence of each flu strain, and x_{T-1} and y_{T-1} represent time-$(T-1)$ estimates of cases for the season if there is no vaccine. The decision at time $T-1$ is to select a strain (X or Y) or defer the decision again to time $T-2$.

The state variables $\theta_{X,t}$ and $\theta_{Y,t}$ are summaries of the information available at time t to predict the upcoming flu season. The particulars of these summaries depend on the specific data collection and belief updating process. The summaries can, for example, be the historical data (the set of observations) or estimates of parameters of a regression model. The statistics for Y are not used in the predictions for X and vice versa.

The following dynamic program gives the recursive relationship for value:

$$V_t(\theta_{X,t}, \theta_{Y,t}) = \max \begin{cases} x_t(\theta_{X,t}) \min(rt, pz)/z, \\ \qquad \text{commit to } X \text{ at period } t, \\ y_t(\theta_{Y,t}) \min(rt, pz)/z, \\ \qquad \text{commit to } Y \text{ at period } t, \\ E[V_{t-1}(\tilde{\theta}_{X,t-1}, \tilde{\theta}_{Y,t-1}) \mid \theta_{X,t}, \theta_{Y,t}], \\ \qquad \text{defer}, \end{cases} \tag{1}$$

where the expected value in the defer expression is taken over the future summaries $\tilde{\theta}_{X,t-1}$ and $\tilde{\theta}_{Y,t-1}$. The expressions for the commit actions use the following assumptions: We assume that the vaccine is distributed uniformly over the population, it is 100% effective against the flu of the respective strain, and it has no cross-effectiveness against the other strain. In addition, we assume that each case of flu is equally severe. Section 6 examines the implications of relaxing these assumptions. The $\min(rt, pz)$ represents the number of doses administered, the lesser of the number of doses produced, and the number demanded. Of the doses administered, the proportion x_t/z or y_t/z (depending on which strain is selected) of them will result in cases of flu prevented. With no time left at $t = 0$, nothing can be produced, so the boundary condition is

$$V_0(\theta_{X,0}, \theta_{Y,0}) \equiv 0. \tag{2}$$

Note that this boundary condition implies that any vaccine produced after the beginning of the flu season has no value. If fact, vaccine produced after the season begins may have some value. The sensitivity analysis in §6 discusses this hard deadline assumption.

The dynamic programming equation for the optimal value is the maximum of the values of the three possible choices—commit to one definitive alternative or the other, or defer until the next period and then choose optimally. To be as general as possible, we do not specify a functional form for the forecasted cases of flu for the season or make specific assumptions about the updating process; instead, we make structural assumptions about the process that generates the sequence of forecasts. The two assumptions for the X trend are stated explicitly below and analogous statements for Y are also made.

Certainly, there will be changes in the forecast cases of flu from one period to the next contingent on new information. However, we assume that on average, there are not anticipated changes in the forecast. We make this explicit by assuming that the mean forecast for the next period is the current forecast. Formally,

Assumption 1. $E[x_{t-1}(\tilde{\theta}_{X,t-1}) \mid \theta_{X,t}] = x_t(\theta_{X,t})$.

This assumption implies that all available information is incorporated into the current forecast.

The next assumption extends this to say that a higher forecast in the current period implies a "higher distribution" (in the sense of first-order stochastic dominance) in the next period. This assumption about the stochastic process, referred to as the *stochastically increasing property*, can be thought of as a persistence or regularity condition. It does not imply that the process itself is necessarily increasing because random walks and mean-reverting processes both satisfy this property, and neither of those is necessarily increasing. Instead, this assumption implies that good news or bad news now tends to persist in the next period. First-order stochastic dominance is equivalent to a statement about the expected value of increasing functions.

Assumption 2. *For increasing* g,

$E[g(x_{t-1}(\tilde{\theta}_{X,t-1})) \mid \theta_{X,t}, x_t(\theta_{X,t})$

is increasing in $x_t(\theta_{X,t})$.

5. Analysis

In this section, we present structural properties of the optimal solution to the dynamic decision problem (1).

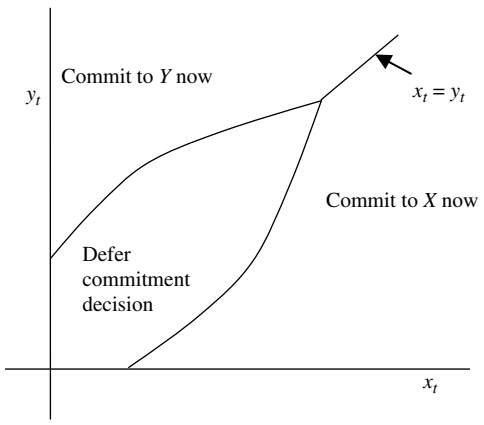
5.1. Threshold Policies

Finding the optimal solution to (1) means that at any time t, we can say which action is optimal: Commit to X (that is, include strain X in the vaccine), commit to Y, or defer the decision until the next period. The form of the optimal solution can be derived by identifying regions in (x_t, y_t) space for which each of the possible actions is optimal. If the vaccine could be produced fast enough so that the total production rt exceeds the number of people who were interested in getting the vaccine pz, then it would make sense to defer until at least just before $t = pz/r$ (which would be small if r was large). While waiting, additional information will become available that may increase the chance of selecting the better strain to include in the vaccine. If, however, $rt < pz$, then further deferral will cause or exacerbate a shortfall in supply.

Figure 3 shows a two-dimensional representation of the optimal solution for a single decision point in the $rt < pz$ case. The solution has a threshold structure that depends on the relative levels of x_t and y_t. If one had to commit to either X or Y, the optimal strategy would be select X if $x_t \geqslant y_t$ and vice versa because we are assuming all flu cases are equally bad. With the defer option possible, some of the areas where X or Y should be chosen otherwise become areas where it is optimal to defer. Specifically, for high enough x_t commit to X: the cutoff for "high enough" is an increasing function of y_t. For high enough y_t, commit to Y: the cutoff for "high enough" is an increasing function of x_t.

Similar to Wald (1947), McCardle (1985), and Kornish (2006), when the information points definitively one way or the other, commit, otherwise, continue observing. This result is stated more formally as follows and is proved in the appendix:

Figure 3. The form of the optimal solution at a given time.



Proposition 1. (a) *At period t, if $r(t-1) \geqslant pz$, the optimal action is to defer.*

(b) *At period t, if $r(t-1) < pz$: For each y_t, there exists a pair of numbers $\xi^L_{X,t}(y_t; \theta_{X,t}, \theta_{Y,t})$ and $\xi^U_{X,t}(y_t; \theta_{X,t}, \theta_{Y,t})$ such that for $x_t \geqslant \xi^U_{X,t}(y_t; \theta_{X,t}, \theta_{Y,t})$, commit to X, for $x_t < \xi^L_{X,t}(y_t; \theta_{X,t}, \theta_{Y,t})$, commit to Y, otherwise, defer the decision. The lower and upper limits $\xi^L_{X,t}(y_t; \theta_{X,t}, \theta_{Y,t})$ and $\xi^U_{X,t}(y_t; \theta_{X,t}, \theta_{Y,t})$ are increasing in y_t.*

Note that $\xi^L_{X,t}(y_t; \theta_{X,t}, \theta_{Y,t})$ may be zero. In other words, for a given y_t, there may no values of x_t for which it is optimal to commit to Y.

5.2. Value of Deferring the Decision

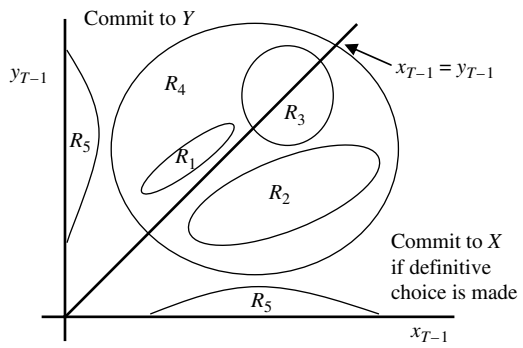
As stated in Proposition 1 and as shown in Figure 3, if the forecasts for the two strains are close and relatively low, then deferring will be the best choice. To gain an intuitive appreciation for where the value of deferring the decision comes from, we will describe several scenarios and explain the relative attractiveness of deferring and acting now.

Consider Figure 4, which illustrates several different regions that indicate the possibilities for future forecasts of both strains of flu. The figure shows the region of nonzero probability for the future forecasts x_{T-1} and y_{T-1}, with these possibilities estimated at time T. The figure shows five scenarios, or regions, representing five different states of information. In each scenario, the assessment of the mean number of flu cases for each strain must ultimately fall within the region for that scenario. Finally, assume in the figure that $rt < pz$ for all of the regions.

If a commitment to a strain must be made in period T, then the strain corresponding to the larger of x_T and y_T should be chosen. The value of deferral comes from the fact that new information may lead to switching the decision from strain X at period T to strain Y at period $T-1$ or vice versa.

If region R_1 represents the set of possible future forecasts given the information known at time T, then there is relatively little uncertainty about the ultimate number of flu

Figure 4. Five information scenarios (regions of possible (x_{T-1}, y_{T-1})) viewed from time T to illustrate the value of deferral.



cases and, more importantly, $y_T > x_T$ in all of the possible resolutions of that uncertainty. Hence, the best decision at time T is to produce the vaccine with the Y strain because this choice will prove to have a higher value for all possible future scenarios. With region R_2, even though there is more relative uncertainty, it will always turn out better if strain X is chosen for the vaccine because $x_T > y_T$ for all resolutions. Hence, commit to strain X now.

In the situation with the region R_3, suppose that x_T is slightly larger than y_T. As indicated, it may ultimately turn out that $y_T > x_T$ so that selection strain X would prove to be inferior to selecting Y. However, because x_T and y_T will always be close given region R_3, the implication of the foregone vaccine production due to deferring the decision at time T would likely be more significant than any benefit due to selecting the strain that would ultimately be chosen given even perfect information at time $T-1$.

Region R_4 has a great deal of uncertainty at time T. If x_T and y_T are near equal, deferring the decision may be sensible. The benefit of a better informed definitive choice at time $T-1$ might outweigh the cost of foregone vaccine production due to deferral. The case represented by region R_5, which is discontinuous in the figure, arguably poses the most difficult decision problem, and is the case in which deferral has the greatest benefit. Here, at time T, you may be relatively clear on the number of flu cases, but uncertain about which strain will be overwhelmingly dominant this year. Suppose that at the current time T, you feel the forecasts x_T and y_T are about equal. Selecting a vaccine at T would result in a half chance of no value because you had the "wrong" strain. Deferring the decision may lead to clarity on which strain will dominate. At the extreme, the measured value is equal to a half chance of avoiding all of the anticipated cases minus the sure value loss associated with lost production by deferring.

For R_3 and R_4, if a choice is made at T, the chance that this choice would have been deemed "wrong" at $T-1$ is about 50%. However, if the wrong strain is chosen given R_3, the discrepancy of flu cases avoided between the right and wrong strains will not be too great. Under R_5, the chance of being wrong could be the same, but the consequences of being wrong are much more significant. In some years, the VRBPAC does not have much deliberation over strain selection: those years represent the easy strain decisions, captured by the R_1 and R_2 scenarios. News coverage about the vaccine composition decision is higher in the years that require substantial deliberation, represented by R_5, because the consequences of the wrong strain choice are higher.

5.3. The Effect of the Deadline on the Thresholds

The optimal thresholds described in Proposition 1 balance the cost of deferring with the benefit of deferring described in §5.2. This trade-off between costs and benefits of deferring depends on how far away the deadline is, and as the deadline approaches, the thresholds change so that the deferral region shrinks. In the extreme: in the final decision period (i.e., when $t = 1$), the deferral region disappears—the optimal action is to commit to X or Y because $V_0 = 0$ (from (2)); deferring in the final period results in no cases of flu prevented.

There are two forces that shrink the deferral region as time passes. The first force is the diminishing size of the stakes. The value of each commit alternative is the product of the forecast for the strain and the number of doses of the vaccine administered. If $rt < pz$, then the value of each commit alternative is proportional to t. When t is high, there is a higher premium on picking the more prevalent strain, and the way to be surer of picking the more prevalent strain is to defer the decision and observe another period of data. The second force shrinking the deferral region is the effect of the information gathering process on the distributions around future forecasts. As we explained in §5.2, a more diffuse distribution over future forecasts increases the value of deferring. If the information gathering process is such that the distribution over future forecasts is contracting with additional information, then that characteristic also makes the deferral region shrink.

We formalize this notion of an information gathering process that tightens the distributions with the following assumption, using the "increased riskiness" concept and result (the result that a mean-preserving spread on the distribution of a random variable increases the expected value of any convex function of the random variables) of Rothschild and Stiglitz (1970).

ASSUMPTION 3. *For convex g,*

$$E[g(x_t(\tilde{\theta}_{X,t})) \mid \theta_{X,t+1}, x_{t+1}(\theta_{X,t+1})] \geqslant E[g(x_{t-1}(\tilde{\theta}_{X,t-1})) \mid \theta_{X,t}, x_t(\theta_{X,t})]$$

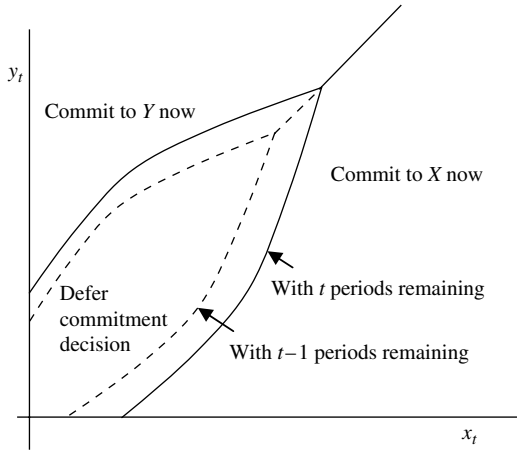
for $x_t = x_{t+1}$.

The assumption states that as the deadline approaches, the distribution on the forecast tightens. In other words, distributions on forecasts at time t based on time $t+1$ information are more spread out (in the sense of the mean-preserving spread) than distributions on forecasts at time $t-1$ based on time t information. A stationary Bayesian Normal-Normal process (see, e.g., Raiffa and Schlaifer 1961) has this property. Adding Assumption 3 to the two earlier assumptions gives us the following result, which is proved in the appendix.

PROPOSITION 2. *If the optimal strategy at time t is to defer, based on information $\theta_{X,t}, \theta_{Y,t}$, with current forecasts x_t and y_t, then with the same forecasts in the earlier period at time $t+1$ (i.e., $x_{t+1} = x_t$ and $y_{t+1} = y_t$), and an information gathering process that is distribution tightening (in the sense of Assumption 3), the optimal strategy is to defer.*

To understand why the optimal strategy is influenced by the deadline in this way, we look at the costs and benefits of deferring as time passes. As described in §5.2, the

Figure 5. The optimal deferral region is smaller with less time remaining until the deadline.



benefit of deferring comes from the possibility that a different commit alternative will be found to be better than the currently indicated choice and will result in a reduction in future flu cases. As the deadline approaches, there is less time for production to garner the benefits. Therefore, for decision periods of equal lengths, the costs of deferring remain constant, but the benefits of deferring decrease over time. Hence, the optimal deferral region should decrease over time, as shown in Figure 5.

To illustrate how Proposition 2 affects the optimal decisions in the flu problem, consider the decision at time T in Figure 2 and assume that a definitive choice of strain will be made at time $T-1$ at the latest and assume that $rT < pz$. The difference in value between deferring and acting can be decomposed into the benefit of deferring and the cost of deferring. For illustration, assume that $x_T > y_T$:

$$\begin{aligned}
&\text{value of deferring} - \text{value of committing now}\\
&= E\max\{x_{T-1}r(T-1)/z, y_{T-1}r(T-1)/z\} - x_T rT/z\\
&= \text{benefit of deferring} - \text{cost of deferring}\\
&= \underbrace{E\max\{x_{T-1}r(T-1)/z, y_{T-1}r(T-1)/z\} - x_T r(T-1)/z}_{\text{benefit of deferring}}\\
&\quad - \underbrace{x_T r(1/z)}_{\text{cost}}\\
&= \underbrace{r(T-1)/z[E\max\{x_{T-1}, y_{T-1}\} - x_T]}_{\text{benefit}} - \underbrace{x_T r(1/z)}_{\text{cost}}.
\end{aligned}$$

Note that the cost of deferring is constant, but the benefit of waiting decreases as time until the deadline decreases: the benefit of deferring is proportional to $(T-1)$ and the coefficient of proportionality is positive because $E(x_{T-1}) = x_T$ so the $E(\max(x_{T-1}, y_{T-1}) - x_T)$ is nonnegative. When there is more time until the deadline, more doses of vaccine can be produced, so choosing between X and Y has more of an effect on value.

5.4. Rate of Production

In the 2003 deliberations for the flu vaccine, some members of the VRBPAC expressed dismay that the vaccine recommendation must come so early in the year (VRBPAC 2003). One natural question to ask is: What is the effect of increasing the production rate r? Clearly, this will in general increase the expected number of flu cases avoided because the decision to commit to a strain in the virus can be made closer to the flu season with no penalty. But how would such a change affect the decisions?

If r can be increased, but t is small enough so that $rt < pz$, what effect does that have on the optimal strategy? Does an increase in r make deferring *more* attractive because production is faster in the remaining time, or does it make deferring *less* attractive because more production is sacrificed while you wait?

The essential issue in this analysis is the relative change in the attractiveness of the commit alternatives, not the comparison of alternatives across different levels of r. Interestingly, in the case in which time is a binding constraint (i.e., more people would get vaccinated if more doses of the vaccine were available) and the production rates r are the same for both strains, we find the following result (which relies only on Assumptions 1 and 2):

PROPOSITION 3. *If the production rate for the vaccine of each strain is equal and increases to r_2 from its original value of r_1, then*

(a) *at period t, if $r_2(t-1) \geqslant pz$ (i.e., $t \geqslant (pz/r_2)+1$), the optimal strategy is to defer, and*

(b) *at period t, if $r_2 t < pz$ (i.e., $t < pz/r_2$), the optimal strategy is the same as it was for r_1.*

This proposition describes the optimal strategy for all decision periods except one. For a period that begins after $(pz/r_2)+1$ but before pz/r_2, we cannot conclude that deferral is optimal nor that the strategy under r_1 is optimal. As a practical matter, the length of the period can be selected to balance precision and computation. It follows from this result, proved in the appendix, that the optimal strategy may change for any decision made between $t_1 = (pz/r_1)+1$ and $t_2 = (pz/r_2)+1$. Specifically, any decision not to defer given r_1, would switch to defer with r_2. For any decision made prior to $t_1 = (pz/r_1)+1$, the optimal decision is to defer given either production rate r_1 or r_2. For decisions made after $t_2 = pz/r_2$, the optimal decisions given either r_1 or r_2 would be the same. Therefore, increasing the production rate from r_1 to r_2 does not change the optimal strategy when time is a binding constraint.

The intuition for this result is as follows. In the $rt < pz$ case, the values of both commit alternatives are proportional to r, and essentially the r "cancels out" in the comparisons. In addition, such a change does not affect the balance between the commit strategies and the defer strategy because ultimately the defer strategy also depends on the relative attractiveness of the two strains. This logic points to a more general result: if the commit alternatives

are both proportional to the same parameters, then changes in that parameter do not affect the optimal decisions. Therefore, similar logic would apply to an analysis of change in vaccine effectiveness assuming symmetry.

Proposition 3 does not hold for the case where vaccines with strains X and Y can be produced at different rates. If instead of a single r, there is a production rate r_X for X and a production rate r_Y for Y, improvement in only one of the rates increases the value of producing the corresponding vaccine relative to both of the other alternatives.

PROPOSITION 4. *For $r_X t < pz$ and $r_Y t < pz$, an increase in r_X lowers the thresholds $\xi^L_{X,t}(y_t; \theta_{X,t}, \theta_{Y,t})$ and $\xi^U_{X,t}(y_t; \theta_{X,t}, \theta_{Y,t})$.*

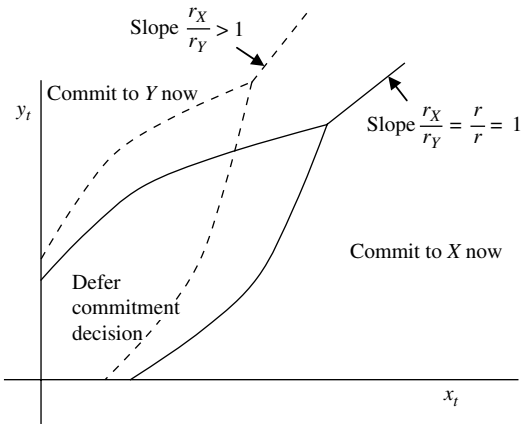
This proposition, proved in the appendix, says that the vaccine strain whose production rate increases gains territory (i.e., regions of (x_t, y_t)) where it is the optimum, the other strain loses territory, and the deferral region can either gain or lose on net. Figure 6 illustrates that for some states, the optimal action changes from defer to commit to vaccine X; for some others, the optimal action changes from commit to Y to commit to X, and for still others, the optimal action changes from commit to Y to defer. This shift happens because increasing the rate of production of vaccine X increases the value of the commit to vaccine X strategy the most, then the value of the defer strategy, and does not affect the value of the commit to vaccine Y strategy.

6. Use of the Model for the Vaccine Composition Decision

As evident from the transcripts of the VRBPAC deliberations in recent years (VRBPAC 2003, 2004, 2005), the committee does not seem to have a systematic procedure to analyze the vaccine composition decision's commit-or-defer structure. Our model provides the logic and a method to do this balancing act.

As with any model, the vaccine composition decision model necessarily includes assumptions about the model structure, information availability and uncertainties, the objective function, and the parameters in the model. In making assumptions, we had four partially conflicting goals: to reflect the reality of the decision problem, to provide insight for the influenza vaccine selection decision, to allow the model to be analyzed, and to maintain generalizability. In this section, we examine what happens to the model analysis if our various assumptions are changed. Hence, this section is essentially a qualitative sensitivity analysis that examines the robustness of the insights from the model.

Figure 6. Shifts in the optimal regions with an increase in r_X.



6.1. Appraisal of the Basic Model

The basic model addresses a vaccine for only one category of virus and yet the vaccines include protection against three viruses. Because the trivalent vaccine is the composition of three reasonably separate decisions as described in §3, focusing on one category of the virus is an appropriate level to lend useful insight and yet simplify the analysis.

While we have presented the model and results as a choice between two strains for mathematical, graphical, and expositional simplicity, many of the results and most of the intuitions follow directly to a choice among $n > 2$ strains. One important difference in the more general case is that the threshold for any particular strain now depends on information about all the other strains. Given that change in Proposition 1, the other three propositions retain their spirits: continuation regions shrink as the deadline looms; for some values of t, decisions are invariant to symmetric changes in production rate; and decisions do change for asymmetric changes in production rate by expanding the commit region for the strain favored by the change.

Another assumption in the basic model is that VRBPAC faces a choice between specific and known strains, as opposed to a choice between currently identified strains and some as yet unknown but emerging strains. Because of the general way in which we have set up the information structure in this model (i.e., the θs and Assumptions 1–3), the model does allow for an unspecified but emerging strain by using subjective assessments of the forecast distribution and the updating process. However, we do not model an unlimited number of new strains, as in the "search for new offers" models (in which new job wages or new prices are repeatedly drawn from a distribution) described in §2. As a practical matter, if the strain is not even identified at the time of deliberation, then the preparation work such as creating a reference strain and establishing the manufacturability of the strain could not be done in time for the coming season.

By considering each category in isolation, the basic model does not allow us to consider strategies such as producing vaccines for two virus categories beginning at one date and deferring production of a vaccine for a third category. Such strategies may be important and could be considered in another model for which our model may provide a useful starting point.

Another important assumption in this model is the fixed and known deadline. Clearly, units of vaccine that are produced during the season may also have some use. For example, in the 2004–2005 flu season, with Chiron's late removal from the supply chain, the government appealed to the other firms to ramp up in-season production. While disruptions occur, the producers strive to have their doses ready for distribution by the start of the season to give full coverage for the season. To examine the benefits of planning to produce throughout the season, we would need to model the demand for vaccine as the season progressed: of the original population, who was interested in vaccination at the beginning of season, and how many would still be interested in December?

It is worthwhile to point out that the hard deadline assumption may be less appropriate if we developed vaccines that could be produced much more quickly than the current vaccine. In such a case, deferring the start of production for one month would both result in accumulating more relevant information and in a shorter time to make up a fixed shortage of vaccine with production after the deadline.

The basic model assumes that once a strain is chosen, one does not switch. Without including it explicitly, this stopping problem assumes that the switching costs and switching time between producing X and producing Y are prohibitively high, so switching is not considered. The switching time includes preparation of the new virus strains for manufacture. In the VRBPAC meeting of February 16, 2005, the industry representative describes this process as taking about four weeks (VRBPAC 2005, p. 175).

The model can be adapted to account for certain changes in the objective function. Obviously, deaths from influenza are much more significant than cases of the flu that do not result in death. The results of the analysis would directly hold if the fatality rates were the same for flu cases caused by different strains. If flu from one virus was more deadly, then the objective could be a weighted average of cases of the flu and flu fatalities. This would essentially result in an "equivalent number of flu cases" that could be used in the analysis. The same technique could be used to address side effects of vaccinations and even economic costs of the program. With costs, the current model de facto assumes that the economic cost of producing and administering the vaccine would be the same with a vaccine of either strain.

We have assumed that the production rate r is known and constant. In reality, the uncertainty about this rate is an important issue in the deliberations about vaccine composition. In the spirit of Propositions 3 and 4, we can identify conditions under which the uncertainty on the rate will not affect the decisions. If the time until the season is clearly such that $t < pz/r$ for any reasonable value of r, then we can use mean rates of production in the analysis if learning about the rate only occurs when production begins, i.e., once a commitment to one strain or another is made. However, if r might be such that $t > pz/r$, then we must include the uncertainty explicitly because the expressions in the objective function are not linear in r. In this case, the logic behind Proposition 1(a), in which there is zero cost to deferring with sufficient time to produce for those who want to be vaccinated, fails. With uncertain r, there may be a nonzero cost to waiting even if the mean rate of production will yield enough doses. Finally, if learning about the rate can happen in the deferral period, then we also must include the uncertainty explicitly because that modification increases the attractiveness of the defer option.

6.2. Insight for the Current Vaccine Decision

The analysis provides a basis to examine how considerations not explicitly included in the model, but relevant to the problem, should affect the decision. These considerations include the efficacy of the vaccines, the severity of the flu due to different strains, and any cross-effectiveness of the vaccines. Insights about the relevance of these considerations follow a logic similar to that concerning the production rate of vaccines analyzed in Proposition 3. Suppose that the efficacy of a vaccine with either strain X or strain Y is less than the 100% assumed in Proposition 1. If the resulting efficacy is the same for each strain, the effect on the values of all strategies is to reduce their value proportionally. Hence, all the same optimal decisions apply. If, however, the efficacy of a vaccine with strain X will be greater than the efficacy of a vaccine with strain Y, the optimal decision regions in Figure 3 will change to favor including strain X in the vaccine. The effect would be similar in nature to increasing the rate of vaccine production for a vaccine with strain X as illustrated in Figure 6.

The same logic applies if the severity of the flu changes. An increase in the severity of flu caused by strain X means that such flu cases either last longer or have more ill effects. If the anticipated severity of flu due to either strain X or strain Y increases by the same percent, the optimal decisions from Proposition 1 do not change as the consequences of all alternatives are proportionally affected. If we relax the assumption from the basic model that the strains are equally severe, and the anticipated severity of flu due to strain X increases more than the anticipated severity due to strain Y, then the optimal decision regions again move to favor including strain X.

With cross-effectiveness of vaccines, meaning, for example, that a vaccine with strain X will prevent or mitigate some cases of a flu due to strain Y, the results are different. Clearly, using logic similar to that above, if a vaccine with strain X prevents some cases of flu due to strain Y, but a vaccine with strain Y does not prevent any cases of flu due to strain X, this favors selecting strain X and the region for its optimal choice in Figure 3 increases. If both strains X and Y in the vaccine guard equivalently against the other strain, their regions for choosing either strain now increase and the deferral region decreases in Figure 3. To understand this, consider the extreme. If either strain in the vaccine guarded against all cases of either flu, it is clear

that one should never defer vaccine production when the production time constraint is binding, and, in this case, you should not care which strain is included in the vaccine.

Analysis of the model indicates the importance of pursuing increases in vaccine production rates because this would allow the decision about what strain to include in the vaccine to be made later in the year when more information is available. Many in the vaccine community would like to see a culture-derived rather than an egg-derived process for flu vaccine production to increase the production rate. The U.S. government made some large grants to pharmaceutical companies in 2006 to help them develop a faster method of flu vaccine development for avian flu (CNN.com 2006). Absent new technology, the government could explore a variety of hedging strategies, such as paying firms to begin production of both strains while deliberation continues or expanding the database of reference or seed strains (Marwick 1993) to reduce the ramp-up time when a new strain is chosen. These types of strategy would counter the findings in the recent IOM report on the market for vaccines mentioned earlier that suggest that the government's large role in the vaccine markets hinders innovation because of the real or perceived interference with profit.

6.3. Insight for Data Collection

There is a variety of information that is and could be collected about the annual flu season. Our model helps indicate what information is most important. That information, simply stated, is the information most likely to change the decisions about (1) whether to defer the decision about what strain to include in a vaccine, and (2) what strain to include if the choice is made now. As our model and the discussion above indicate, if rates of production, vaccine efficacy, and flu severity are the same for different strains, better estimates of these factors should not affect the vaccine composition decision. Differences in such factors do matter, and so information that might indicate any such differences would be useful to obtain.

The key component to the vaccine composition decision is which strain will cause the dominant number of cases in the coming season. Hence, information collection to forecast flu cases due to different strains and the uncertainty about these forecasts is very important. This uncertainty is an essential feature of the problem, and therefore the forecasting mindset must be to emphasize the construction of probability distributions over numbers of flu cases and not just point estimates. It is the uncertainty in the forecasts that makes the deferring alternative viable and interesting.

To support these forecasting models, representative data needs to be collected. The Centers for Disease Control and Prevention (CDC) and the World Health Organization (WHO) currently collect three relevant streams of data about the spread of the different strains of the virus:

(1) Deaths due to influenza and pneumonia in 122 cities, reported as the percentage of total deaths, so the number can be extrapolated to the entire population.

(2) The percentage of people who come to the doctor each week because they think they have the flu, as reported by a network of hundreds of physicians during flu season.

(3) Laboratory samples sent in for virus-typing.

All three of these streams have issues concerning use and usefulness. Is extrapolating to the population level from the first stream of data appropriate given that the data reported is from urban centers? For the second stream, there could be many reasons for an increase in office visits, such as news coverage of flu deaths. Finally, the third stream suffers from a selection bias: Why are the physicians sending these particular samples in for typing? There are other questions about the data, such as how to use the data being produced in the off-season (e.g., the month of March in the northern hemisphere) to make predictions for the coming season. To what extent can the cross-hemispheric data be used to predict spread for the coming season? The key point is the following: one should ask what information would be most useful and then do the best at gathering it rather than just rely on what data is collected by others because it might not be too useful.

6.4. Insight for Expanding the Vaccine Composition Decision

In addition to the model and data recommendations, the model suggests changes in the flu vaccine composition problem that are worthwhile to consider and analyze.

In our analysis, we looked at the appropriate strain to include in one category (i.e., A1, A2, or B) in isolation. The three strains complicate the decision because while one considers waiting for new data to resolve the A1 forecast, for example, production time may be lost on the vaccines for the other strain categories. The third strain might be separately introduced and added to or used separately from the two-strain production. Alternatively, perhaps after the choice of a third strain, it could be added only to vaccine subsequently produced.

A natural question to ask about the trivalent formulation is: "Why three?" Why not four (include both strains from a category) or two (omit one of the categories)? A variation on this question is: "Why one in each of the three categories?" A vaccine formulation with more than three strains or without the rigid categories changes the "X or Y" problem framing to allow for the "X and Y" possibility. A four-or-more-strain vaccine entails its own balancing acts with the increased demands on production capacity and increased demands on the human body recipients, such as the body's difficulty in handling the increased protein load associated with the expanded dose of the vaccine. Elimination of the rigid categories would be especially important if technological advances allowed the cross-effectiveness across categories to be significant. Choosing the best collection of strains is quite a different problem from choosing the best single strain, akin to a portfolio problem. As in a wide variety of stopping problems, the result we would look for would be that strong evidence suggests a commitment; weak evidence suggests deferment.

7. Relevance of the Model to Other Commit-or-Defer Decisions

In this paper, we have analyzed an important policy issue as an example of a situation with two salient features: a looming deadline and the ability to defer the decision to gather information; in this case, information about the spread of a disease. It is clear that while deferring surely has a cost—lost production—it also has a benefit—the opportunity to make a more-informed decision.

This policy decision has analogs in other decision-making realms because deadlines arise in many contexts. For example, deadlines can come from the holiday shopping season or a car model year. They also arise from events with a particular date, such as the Super Bowl or an industry trade show; from anticipated competitors' actions; from patent expiration or expected regulatory change; and from production scheduling constraints in which components must be assembled. For a wide range of commit-or-defer problems, we would expect to see a solution that advises that if there is strong enough evidence in one direction or the other, commit to that direction. Otherwise, in a "middle region" (literally and figuratively), defer the decision and continue to gather information.

Another insight from the vaccine context for these other realms is to raise the question as to whether it is imperative to make a definitive either-or choice between two options, or whether it might make sense to pursue multiple avenues simultaneously. The vaccine composition decision is currently treated as strictly either-or within each category of the virus. However, for an electronics manufacturer preparing for the annual Consumer Electronics Show, it may make sense to invest in both of two competing standards (e.g., the two current next generation DVD standards, HD and Blu-Ray). In this vein, we suspect that Super Bowl souvenir vendors do not bet the house on only one of the two teams!

The result in Proposition 3 is an example in which optimal decisions are insensitive to symmetric changes in a parameter. This result generalizes to the insight that in deciding between commitments (definitive courses of action) and the option to defer, if a change does not affect the relative values of the definitive courses of action, then it may not need serious study. For example, if a fast-food chain is deciding on children's entertainment promotional tie-ins for the next season, it may not need to consider broader trends of movie attendance versus other forms of entertainment if the leading contenders are both movie characters.

We have used this public sector problem to motivate an analysis that is at the heart of challenging and important issues in a variety of investment problems. Whenever there is something new, we cannot know if it is an important change for the future or a flash in the pan. Is this one of Christensen's (1997) "disruptive technologies" or will it be a short-lived fad? Instead of trying to make a prediction about winners and losers, we prefer to think about these types of problems as sequential decision problems under uncertainty in which investments can be deferred.

8. Conclusions

Many analyses are used to support the decisions and deliberations of the VRBPAC. Data are collected and analyzed: data about virus spread, production techniques, and scientific opinions about the evolution of the viruses. Waiting for more information has been discussed, but not analyzed as a sequential decision problem under uncertainty. However, one hurdle to implementation is that the forecasting models and the data to populate them are not readily available. Development of both of those elements would take significant effort.

The flu epidemics are notoriously difficult to predict, and like many applications of decision analysis, difficult predictions, which often result in distributions with very broad spreads, are not readily welcomed by experts, who are accustomed to sharing what they know, not the degree to which they do not know something. However, it is this very aspect itself—the difficulty of knowing and the broad range of possibilities—that makes waiting for more information an attractive alternative.

Recognizing these realities of the flu vaccine composition decision, our intent was to develop and analyze a model that would provide useful insights for this significant and complex annual decision: The model is general in order to preclude the requirement of information beyond that which could possibly be obtained. It allows for numerous what-if analyses to examine the potential effects on decisions of both directional and magnitude changes of relevant problem features such as vaccine production rates, dominant strains causing the flu, the numbers of flu cases, severity of the flu, and effectiveness of the vaccines. The numerous sensitivity analyses of assumptions and parameters in the model provide several insights for improved communication among those responsible for selecting flu vaccines. The intent is to support better informed choices, which hopefully leads to better decisions and lessened impact of the dreaded flu season.

Appendix

Proof of Proposition 1

Part 1a. Show that if t is such that $r(t-1) \geqslant pz$, then the optimal action is to defer. Define $t^* \equiv pz/r$, the point in time at which there is exactly enough time left to produce a dose of vaccine for everyone who wants one. If $t-1 > t^*$, then at time t, one can produce enough doses to meet demand, and by deferring one period, one can still produce enough doses to meet demand. Then, there is a benefit to waiting (in better information), but no cost (in doses administered).

Part 1b. The case $r(t-1) < pz$ is made of two subcases: $rt < pz$ and $rt \geqslant pz$. First, we treat the case that $rt < pz$. We start with the upper bound. We show that for a given y_t, there is a cutoff $\xi^U_{X,t}(y_t; \theta_{X,t}, \theta_{Y,t})$ such that for $x_t \geqslant \xi^U_{X,t}(y_t; \theta_{X,t}, \theta_{Y,t})$, it is optimal to commit to X now. We do this by showing that once x_t is at a level such that committing to X is optimal, then committing to X is optimal for any larger x_t. It suffices to show that $V_t(\theta_{X,t}, \theta_{Y,t}) - x_t rt/z$ is decreasing in x_t. Use induction on the number of periods left:

(1) First, show that it holds for $t = 1$. It does because $\max\{x_1 r/z, y_1 r/z, 0\} - x_1 r/z$ is decreasing in x_1.

(2) Now assume that the claim is true for $t-1$: $V_{t-1}(\theta_{X,t-1}, \theta_{Y,t-1}) - x_{t-1} r(t-1)/z$ is decreasing in x_{t-1}.

(3) Show that it follows that the claim holds for t. Show that $\max\{x_t rt/z, y_t rt/z, E[V_{t-1}(\theta_{X,t-1}, \theta_{Y,t-1}) \mid \theta_{X,t}, \theta_{Y,t}]\} - x_t rt/z$ is decreasing in x_t.

If either of the first two arguments of the max expression yields the maximum, the claim obviously holds. The third argument needs further attention; show that $E[V_{t-1}(\theta_{X,t-1}, \theta_{Y,t-1}) \mid \theta_{X,t}, \theta_{Y,t}] - x_t rt/z$ is decreasing in x_t.

Rewrite that expression by adding and subtracting $E[x_{t-1} r(t-1)/z \mid \theta_{X,t}, \theta_{Y,t}]$ and then rearranging to get $E[V_{t-1}(\theta_{X,t-1}, \theta_{Y,t-1}) - x_{t-1} r(t-1)/z \mid \theta_{X,t}, \theta_{Y,t}] + E[x_{t-1} r(t-1)/z \mid \theta_{X,t}, \theta_{Y,t}] - x_t rt/z$.

The expression inside the first expectation is decreasing in x_{t-1} by the induction hypothesis. Using Assumption 2, we conclude that the first term (i.e., the first expectation term) is decreasing in x_t. Now it suffices to show that $E[x_{t-1} r(t-1)/z \mid \theta_{X,t}, \theta_{Y,t}] - x_t rt/z$ is decreasing in x_t. By Assumption 1, the first term is equivalent to $x_t \cdot r(t-1)/z$, so the difference is $-x_t r/z$, which is decreasing in x_t.

Now, we address the lower bound: For a given y_t, there is a cutoff $\xi^L_{X,t}(y_t; \theta_{X,t}, \theta_{Y,t})$ such that for $x_t < \xi^L_{X,t}(y_t; \theta_{X,t}, \theta_{Y,t})$, it is optimal to commit to Y now. We do this by showing that if x_t is such that committing to Y is optimal, then committing to Y is optimal for any smaller x_t. It suffices to show that $V_t(\theta_{X,t}, \theta_{Y,t}) - y_t rt/z$ is increasing in x_t. This is equivalent to showing that $V_t(\theta_{X,t}, \theta_{Y,t})$ is increasing in x_t. That can be shown by induction: with $t = 1$, the value function is the maximum of two expressions, one linearly increasing in x_1 and the other constant. The key step in the induction uses Assumption 2: that the stochastic process is such that the expected value of an increasing function of a random variable is increasing in the random variable.

To show that the threshold functions are increasing, i.e., $\xi^U_{X,t}(y_t; \theta_{X,t}, \theta_{Y,t})$ and $\xi^L_{X,t}(y_t; \theta_{X,t}, \theta_{Y,t})$ are increasing in y_t, we look at what determines the thresholds. The threshold $\xi^U_{X,t}(y_t; \theta_{X,t}, \theta_{Y,t})$ is the maximum of the two x_ts at the intersections of

$$x_t rt/z = y_t rt/z \quad \text{and}$$
$$x_t rt/z = E[V_{t-1}(\theta_{X,t-1}, \theta_{Y,t-1}) \mid \theta_{X,t}, \theta_{Y,t}].$$

The x_t solving the first equation is clearly increasing in y_t. The x_t solving the second equation is also increasing in y_t by Assumption 2 and the fact that V_t is increasing in x_t and y_t. Therefore, if both x_t points are increasing in y_t, then the maximum of the two points is increasing in y_t.

Likewise, the threshold $\xi^L_{X,t}(y_t; \theta_{X,t}, \theta_{Y,t})$ is the minimum of the two x_ts at the intersections of

$$y_t rt/z = x_t rt/z \quad \text{and}$$
$$y_t rt/z = E[V_{t-1}(\theta_{X,t-1}, \theta_{Y,t-1}) \mid \theta_{X,t}, \theta_{Y,t}].$$

The x_t solving the first equation is clearly increasing in y_t. The x_t solving the second equation is also increasing in y_t: both sides of the equation are increasing in y_t, but the left-hand side (LHS) increases more than the right-hand side (RHS) for a given change in y_t. The LHS increases by rt/z for a unit change in y_t, and the RHS increases by at most $r(t-1)/z$ for a unit change in y_t. Therefore, if both x_t points are increasing in y_t, then the minimum of the two points is increasing in y_t.

Finally, we return to the case that $r(t-1) < pz$, but $rt \geqslant pz$. At time t, there is sufficient time to meet all of demand, but at the start of the subsequent period, there would not be sufficient time. All of the arguments above continue to hold, except that at the start of period t (and that period only), the "commit" alternatives have values of $x_t p$ and $y_t p$ (instead of $x_t rt/z$ and $y_t rt/z$).

Proof of Proposition 2

At time t, pick a point in (x_t, y_t) space such that it is optimal to defer. That is,

$$x_t rt/z \leqslant E[V_{t-1}(\theta_{X,t-1}, \theta_{Y,t-1}) \mid \theta_{X,t}, \theta_{Y,t}] \quad \text{and}$$
$$y_t rt/z \leqslant E[V_{t-1}(\theta_{X,t-1}, \theta_{Y,t-1}) \mid \theta_{X,t}, \theta_{Y,t}].$$

To show that the optimal strategy is to defer at time $t+1$, we must establish that

$$x_{t+1} r(t+1)/z \leqslant E[V_t(\theta_{X,t,Y,t}) \mid \theta_{X,t+1}, \theta_{Y,t+1}] \quad \text{and}$$
$$y_{t+1} r(t+1)/z \leqslant E[V_t(\theta_{X,t}, \theta_{Y,t}) \mid \theta_{X,t+1}, \theta_{Y,t+1}].$$

From the condition in the proposition, $x_{t+1} = x_t$ and $y_{t+1} = y_t$, so it suffices to show that

$$E[V_t(\theta_{X,t}, \theta_{Y,t}) \mid \theta_{X,t+1}, \theta_{Y,t+1}] - E[V_{t-1}(\theta_{X,t-1}, \theta_{Y,t-1}) \mid \theta_{X,t}, \theta_{Y,t}] \geqslant x_t r/z \quad \text{and}$$
$$E[V_t(\theta_{X,t}, \theta_{Y,t}) \mid \theta_{X,t+1}, \theta_{Y,t+1}] - E[V_{t-1}(\theta_{X,t-1}, \theta_{Y,t-1}) \mid \theta_{X,t}, \theta_{Y,t}] \geqslant y_t r/z.$$

We cover the proof of the first inequality. The second inequality is analogous. By Assumptions 1 and 3 and the convexity of V_t, it suffices to show that $E[V_t(\theta_{X,t}, \theta_{Y,t}) - V_{t-1}(\theta_{X,t}, \theta_{Y,t}) - x_t r/z \mid \theta_{X,t+1}, \theta_{Y,t+1}] \geqslant 0$. The convexity

of V_t can be established with an induction argument, relying on the fact that the maximum of convex functions is convex and the fact that convexity is preserved through the expectation operator (i.e., positive-weighted sums of convex functions are convex).

The last sufficient condition holds if $V_t(\theta_{X,t}, \theta_{Y,t}) - V_{t-1}(\theta_{X,t}, \theta_{Y,t}) - x_t r/z \geqslant 0$. Treating that inequality as the basis of an induction, we see it holds for $t = 1$ because $\max\{x_1 r/z, y_1 r/z\} - 0 \geqslant x_1 r/z$. It is also true that if the inductive inequality holds for $t-1$, it then holds for t.

By the inductive assumption, for equivalent forecast values, deferring at time $t-1$ implies deferring at time t. Therefore, $V_t(\theta_{X,t}, \theta_{Y,t}) - V_{t-1}(\theta_{X,t}, \theta_{Y,t}) - x_t r/z \geqslant 0$ in all possible cases: deferral at times t and $t-1$ (by the inductive assumption), deferral at time t and commitment at time $t-1$, and commitment at times t and $t-1$.

Proof of Proposition 3

Part 3a. The proof is the same as the proof of Proposition 1, part a.

Part 3b. The proof that changes in r do not affect the optimal decisions for $rt < pz$ is by induction:

(1) Show that the result holds with one period to go, $t = 1$:

$$V_1(\theta_{X,1}, \theta_{Y,1}) = \max\{x_1 r/z, y_1 r/z, 0\} = r \max\{x_1/z, y_1/z, 0\}.$$

The comparison between the three alternatives does not change as r changes, e.g., the set of (x_1, y_1) for which $x_1 r/z \geqslant y_1 r/z$ is not affected by r.

(2) Assume that the result holds with $t-1$ periods remaining. In other words, assume that the three comparisons between the alternatives do not depend on r with $t-1$ periods remaining because the r can be factored out. We define $W_{t-1}(\theta_{X,t-1}, \theta_{Y,t-1})$ as V_{t-1} without the r factor. Define

$$W_{t-1}(\theta_{X,t-1}, \theta_{Y,t-1}) = \max \begin{cases} x_{t-1}(t-1)/z \\ y_{t-1}(t-1)/z \\ E[W_{t-2}(\tilde{\theta}_{X,t-2}, \tilde{\theta}_{Y,t-2}) \mid \theta_{X,t-1}, \theta_{Y,t-1}] \end{cases}$$

and $W_0 \equiv 0$, and assume that $V_{t-1} = rW_{t-1}$.

(3) Show that the result holds with t periods remaining, i.e., show that $V_t = rW_t$,

$$W_t(\theta_{X,t}, \theta_{Y,t}) = \max \begin{cases} x_t t/z, \\ y_t t/z, \\ E[W_{t-1}(\tilde{\theta}_{X,t-1}, \tilde{\theta}_{Y,t-1}) \mid \theta_{X,t}, \theta_{Y,t}], \end{cases}$$

$$rW_t(\theta_{X,t}, \theta_{Y,t}) = \max \begin{cases} x_t rt/z, \\ y_t rt/z, \\ rE[W_{t-1}(\tilde{\theta}_{X,t-1}, \tilde{\theta}_{Y,t-1}) \mid \theta_{X,t}, \theta_{Y,t}]. \end{cases}$$

Because r is a constant, it can be moved inside the expectation operator in the third term above, so that term becomes $E[rW_{t-1}(\tilde{\theta}_{X,t-1}, \tilde{\theta}_{Y,t-1}) \mid \theta_{X,t}, \theta_{Y,t}]$.

By the induction hypothesis, $V_{t-1} = rW_{t-1}$, so we can write

$$rW_t(\theta_{X,t}, \theta_{Y,t}) = \max \begin{cases} x_t rt/z, \\ y_t rt/z, \\ E[V_{t-1}(\tilde{\theta}_{X,t-1}, \tilde{\theta}_{Y,t-1}) \mid \theta_{X,t}, \theta_{Y,t}], \end{cases}$$

which is equal to V_t.

Proof of Proposition 4

First, we show that an increase in the production rate for X, r_X, lowers the threshold on x_t above which commit to X is optimal, $\xi^U_{X,t}(y_t; \theta_{X,t}, \theta_{Y,t})$. With different production rates, for the case $r_X t < pz$ and $r_Y t < pz$, the optimal value function (1) becomes

$$V_t(\theta_{X,t}, \theta_{Y,t}) = \max \begin{cases} x_t r_X t/z, & \text{commit to } X \text{ at period } t, \\ y_t r_Y t/z, & \text{commit to } Y \text{ at period } t, \\ E[V_{t-1}(\tilde{\theta}_{X,t-1}, \tilde{\theta}_{Y,t-1}) \mid \theta_{X,t}, \theta_{Y,t}], & \\ \qquad \text{defer.} & \end{cases}$$

The threshold on x_t above which it is optimal to commit to X is determined by two comparisons:

$$x_t r_X t/z = y_t r_Y t/z \quad \text{and} \tag{3}$$

$$x_t r_X t/z = E[V_{t-1}(\theta_{X,t-1}, \theta_{Y,t-1}) \mid \theta_{X,t}, \theta_{Y,t}]. \tag{4}$$

In comparison (3), the LHS is increasing in x_t and r_X and the RHS is not a function of either. Therefore, an increase in r_X shifts the LHS up, reducing x_t at which the LHS and RHS intersect. In comparison (4), both the LHS and RHS are increasing in x_t and r_X. The LHS has a slope of $r_X t/z$ with respect to x_t. The RHS, the value of deferring, is an expected combination of future commit values, so the slope of the RHS with respect to x_t cannot be more than $r_X(t-1)/z$. (There may be additional terms in the value of deferring, related to y_t, but they will not affect the slope of the RHS with respect to x_t.) Therefore, the slope of the LHS is greater than the slope of the RHS w.r.t. x_t. As r_X increases, both slopes increase. The slope of the LHS changes by t/z for a unit increase in r_X. The slope of the RHS changes by at most $(t-1)/z$, so the point of intersection decreases. The threshold for x_t, above which it is optimal to commit to X is the maximum of x_t that satisfies (3) and x_t that satisfies (4). If both points are lowered, then the maximum is also lowered. An increase in r_X lowers x_t that satisfies both conditions, so the threshold is reduced with an increase in r_X.

Now we show that an increase in the production rate for X, r_X, lowers the threshold on x_t below which commit to Y is optimal, $\xi^L_{X,t}(y_t; \theta_{X,t}, \theta_{Y,t})$. The threshold on x_t below which it is optimal to commit to Y is determined by two comparisons. The first comparison is the same as (3) above. The second comparison is

$$y_t r_Y t/z = E[V_{t-1}(\theta_{X,t-1}, \theta_{Y,t-1}) \mid \theta_{X,t}, \theta_{Y,t}]. \quad (5)$$

Here, the LHS is not a function of x_t or r_X. The RHS is increasing in x_t and r_X. So, an increase in r_X raises the slope of the RHS with respect to x_t, decreasing the point of intersection. The threshold $\xi^L_{X,t}(y_t; \theta_{X,t}, \theta_{Y,t})$ is the minimum of the two x_t that solve (3) and (5). These two x_t are decreasing in r_X, so their minimum is also decreasing in r_X.

Acknowledgments

Jonathan Wiener of the Duke Law School suggested this problem for analysis and provided guidance for the paper. The authors gratefully acknowledge helpful conversations with Sam Katz and Dennis Clements of Duke University Medical Center, Frank Sloan of the Fuqua School of Business and the Center for Health Policy Research and Economics at Duke, Mike Stoto of RAND, and Chris Rasmussen and Elisabeth Neumeier of GlaxoSmithKline. The authors greatly appreciate Soo-Haeng Cho's help with the proof of Proposition 2. Many thanks to Noah Perin for research assistance. The usual disclaimer applies.

References

CDC.gov. 2006. Avian influenza. http://www.cdc.gov/avian.

Chow, Y. S., H. Robbins, D. Siegmund. 1971. *Great Expectations: The Theory of Optimal Stopping*. Houghton Mifflin Company, Boston, MA, 49–52.

Christensen, C. L. 1997. *The Innovator's Dilemma*. Harvard Business School Press, Boston, MA.

CNN.com. 2003. Panel reluctantly backed flu vaccine to FDA. December 15. http://www.cnn.com/2003/HEALTH/12/13/sprj.flu03.vaccine/index.html.

CNN.com. 2004. WHO plans flu pandemic summit. November 1. http://www.cnn.com/2004/HEALTH/10/31/flu.pandemic.ap/index.html.

CNN.com. 2006. Drugmakers get $1 billion for bird flu vaccine. May 5. http://money.cnn.com/2006/05/04/news/companies/novartis/index.html.

DeGroot, M. H. 1970. *Optimal Statistical Decisions*. McGraw-Hill Book Company, New York, 277–278.

Fisher, M. L., A. Raman. 1996. Reducing the cost of demand uncertainty through accurate response to early sales. *Oper. Res.* **44**(1) 87–99.

Fisher, M. L., J. H. Hammond, W. R. Obermeyer, A. Raman. 1994. Making supply meet demand in an uncertain world. *Harvard Bus. Rev.* **72**(3) 83–93.

Girshick, M. A. 1946. Contributions to the theory of sequential analysis. *Ann. Math. Statist.* **17**(3) 282–298.

Institute of Medicine. 2004. *Financing Vaccines in the 21st Century: Assuring Access and Availability*. The National Academies Press, Washington, D.C.

Jensen, R. 1982. Adoption and diffusion of an innovation of uncertain profitability. *J. Econom. Theory* **27**(1) 182–193.

Kornish, L. J. 2006. Technology choice and timing with positive network effects. *Eur. J. Oper. Res.* **173**(1) 268–282.

Lippman, S. A., J. J. McCall. 1976. The economics of job search: A survey: Part 1. *Econom. Inquiry* **14**(2) 155–189.

Lippman, S. A., K. F. McCardle. 1991. Uncertain search: A model of search among technologies of uncertain values. *Management Sci.* **37**(11) 1474–1490.

Marwick, C. 1993. Influenza's distant early warning system. *JAMA* **269**(10) 1223.

McCardle, K. F. 1985. Information acquisition and the adoption of new technology. *Management Sci.* **31**(11) 1372–1389.

Mortensen, D. T. 1986. Job search and labor market analysis. O. Ashenfelter, R. Layard, eds. *Handbook of Labor Economics*, 1st ed., Vol. 2. Elsevier, New York, 860–861.

Neustadt, R. E., H. Fineberg. 1982. *The Epidemic That Never Was: Policy-Making and the Swine Flu Scare*. Random House, Inc., New York.

Raiffa, H., R. Schlaifer. 1961. *Applied Statistical Decision Theory*. MIT Press, Cambridge, MA.

Rothschild, M. 1974. Searching for the lowest price when the distribution of prices is unknown. *J. Political Econom.* **82**(4) 689–711.

Rothschild, M., J. E. Stiglitz. 1970. Increasing risk I: A definition. *J. Econom. Theory* **2**(3) 225–243.

Thompson, W. W., D. K. Shay, E. Weintraub, L. Brammer, C. Bridges, N. Cox, K. Fukuda. 2004. Influenza-associated hospitalizations in the United States. *JAMA* **292**(11) 1333–1340.

Treanor, J. 2004. Weathering the influenza vaccine crisis. *New England J. Medicine* **351**(20) 2037–2040.

Vaccines and Related Biologic Products Advisory Committee (VRBPAC). 2003. Transcript of vaccines and related biological products advisory committee meeting. March 18. http://www.fda.gov/ohrms/dockets/ac/03/transcripts/3941t1.doc.

Vaccines and Related Biologic Products Advisory Committee (VRBPAC). 2004. Transcript of vaccines and related biological products advisory committee meeting. February 18. http://www.fda.gov/ohrms/dockets/ac/04/transcripts/4020t1.htm.

Vaccines and Related Biologic Products Advisory Committee (VRBPAC). 2005. Transcript of vaccines and related biological products advisory committee meeting. February 16. http://www.fda.gov/ohrms/dockets/ac/05/transcripts/2005-4087T1.htm.

Wald, A. 1947. *Sequential Analysis*. Wiley Publications in Statistics, New York.

Whalen, J., B. McKay, S. Lueck. 2004. U.S. flu vaccines face rationing as plant is shut. *Wall Street J.* (October 6) A1.

Wu, J. T., L. M. Wein, A. S. Perelson. 2005. Optimization of influenza vaccine selection. *Oper. Res.* **53**(3) 456–476.