Deliberation with a Deadline:

The Influenza Vaccine Composition Decision*

Laura J. Kornish† Ralph L. Keeney, Jonathan B. Wiener

The Fuqua School of Business, Duke University

The Fuqua School of Business, Duke University

Duke Law School

November 14, 2004

Abstract

We analyze deliberation strategies in a setting with a production deadline. In this dynamic decision problem, there is a choice between two courses of action and a third alternative of delaying the commitment to gather more information about the prevalence or popularity of each action. The deadline affects the cost of information gathering, which in this case is the value that is sacrificed by reducing the interval for production. We illustrate the model with the annual influenza vaccine composition decision: deciding between strains of the virus to include, which must happen in a timely manner to allow time for vaccine production before the flu season begins. Drawing on the public nature of the deliberations surrounding this issue, we suggest ways to improve the decision-making process. We describe the links between this public health decision and a firm's choice in the face of emerging trends.

Key Words: Information Gathering, Deadline, Stochastic Dynamic Programming, Influenza Vaccine Composition, Emerging Trends

*We 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. Many thanks to Noah Perin for research assistance. The usual disclaimer applies.

†kornish@duke.edu

1 Introduction

In many important choices, the decision maker has a set of alternatives to choose from and also the alternative of deferring the decision to learn more. Decision deferral has been studied in the operations research and economics literatures, considering both information acquisition (e.g., Jensen 1982 and McCardle 1985) and search for new alternatives (e.g., Lippman and McCall 1976). These "commit-or-wait" choices are ubiquitous in business, government, and individual decisions.

This paper examines an important class of commit-or-wait problem in which there is a fixed deadline after which further efforts are no longer valuable. With a deadline, the cost of information gathering or the cost of delay includes the value that is sacrificed by reducing the interval for action. As a motivating example, we use a public sector choice, the influenza vaccine composition recommendation to the Food and Drug Administration (FDA). This is a classic example of a problem that has information gathering opportunities about a trend and a production deadline. We discuss how the logic of the model applies more broadly to the detection of trends in other settings.

This work has a decision analysis perspective: defining the decision criteria, using the dynamic nature of the problem in developing alternatives, and describing the important uncertainties in the problem. The stochastic dynamic programming model lends insights into the dynamic decision problem, and we (1) offer general conclusions about the trade-offs between immediate action and deferral, (2) make recommendations for the specific example studied, the flu vaccine composition, and (3) discuss links to other applications in the private sector.

2 Literature

This work is in the vein of search for information, formulated as a sequential decision problem under uncertainty. It is related to work in sequential analysis (see e.g., Girshick 1946 and Wald 1947), which applies to problems in which information is gathered at each step and a decision about whether or not to proceed is made. Wald 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.

In the 1980s, Jensen (1982) and McCardle (1985) looked at the adoption decisions for a tech-

nology of unknown profitability. In Jensen's model, information is costlessly observed each period; in McCardle's there is a fixed cost of observation with the option to stop the observation process altogether. The work of McCardle explicitly discusses the analysis of a single new technology, but implicitly it analyzes a choice between something new and a fallback option, with known value. In the present work, similar to Kornish (2004), we analyze the choice between two new technologies, each of which has new information each period relevant to forecasting ultimate impact. Unlike the current paper, Kornish (2004) features competition with positive feedback between the contenders: the success of one product harms the prospect of the other. A distinguishing feature of the current work is the physical production, which is time consuming, and gives import to the deadline in the problem. The observation may not have an explicit cost, but it does have an implicit cost of leaving less time for production.

We use the flu vaccine composition decision as a motivating example throughout the paper. 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, 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. While the book doesn't contain a mathematical analysis, the tone of the information gathering advice in the lessons learned is consistent with our model. The authors include an appendix of questions to guide vaccine policy decisions, indicating that a key question (to be "asked of every group, regardless of expertise" p. 217) 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.

Similar in spirit to our analytical approach is the work of Wu, Wein, and Perelson (to appear). They develop a theory of antigenic distance and propose an optimization of vaccine selection based on the vaccination history of the population. In contrast, our approach focuses on tracking the spread of the candidate strains in the current season.

The vaccine problem is closely related to other watchful waiting problems, such as decisions related to fashions for a new season (see, e.g., Fisher et al. 1994). 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 units that are produced are allocated. With markets for goods, the beneficiaries are the people

whose demand is satisfied; with vaccines, the beneficiaries are never identified—they are the people who don't contract the illness. We discuss this related trend-tracking problem in section 7.2.2.

3 Example background

This work began with a conversation about the parallels between commit-or-wait problems in the public and private sectors. The influenza vaccine decision is a high profile, important, classic commit-or-wait problem in the public sector affecting millions of people in the U.S. each year. Over 80 million people were vaccinated for the 2003-2004 season, and annual influenza deaths average 36,000. The flu vaccine composition decision made headlines in the spring of 2003 because of the emergence of a new viral strain, the Fujian strain (CNN.com, 2003). With the new strain, there was much discussion about the uncertainty about how widespread it would become and whether or not there was time to develop and produce a vaccine before the start of the flu season (the deadline in the problem) in the fall. Currently, vaccine production is in the news for two reasons. First, contamination at Chiron's Liverpool facility has left the U.S. with half the anticipated vaccine supply (Whalen et al., 2004). Second, the specter of a new strain of avian flu has the WHO concerned about a pandemic (CNN.com, 2004).

The flu vaccine decision is a quintessential example of delay and a deadline: each moment spent deliberating on the front end of the problem leaves less time to act on the decision. Much is made of the time-intensive and therefore constraining nature of the vaccine production process, because the time requirements for production force the decision about vaccine composition to be made so early in the year. The vaccine example is also appealing because of the long tradition of analogies between biologic and market phenomenon. The diffusion of innovations (see, e.g., Rogers 1962, 1995 and Bass 1969) draws on the concept of contagion, using the same mathematics to capture word-of-mouth as that used to describe the spread of disease. So for these characteristics: information gathering about a trend, a deadline, and contagion, the flu vaccine is a rich example to understand a problem of deliberation with a deadline.[1]

Each year in early spring, Vaccines and Related Biologic Products Advisory Committee (VRBPAC) recommends to the FDA which strains of flu virus will be included in the vaccine for the next

[1] And on a practical note, the extensive public record and transcript of the committee's deliberation (VRBPAC 2003, 2004) offers us an opportunity to scrutinize the decision-making process.

winter, to be delivered starting in the fall. The influenza viruses are in a perpetual state of flux, with changes in the virus happening in large and small steps (antigenic "shifts" and "drifts" respectively). Due to these changes, in some years, the VRBPAC recommends a new strain for inclusion. In the U.S. and much of the world, the influenza vaccine is a trivalent vaccine, including one type of each of three categories of virus (labeled A1, A2, and B). The committee meets to recommend which virus in each category should be in the vaccine. Because of the annual time frame, if a new strain is included in the virus, 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.

When the committee meets to discuss the recommendation, there is the 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, and any deferral leaves less time for production before the start of flu season. And because the required production time is a binding constraint on the problem, there is a trade-off between producing more and knowing more.

In the vein of knowing more, by delaying the decision, information can be gathered or observed to inform some of the uncertain aspects of the landscape. 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 across strains. By waiting to make the decision, information about these uncertainties can be learned, allowing for a more-informed decision. 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.

This vaccine composition choice shares some aspects of the technology adoption problem in the face of competing standards studied by Kornish (2004). In both models, delay can be used to try to understand 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 a situation in which the decision maker is observing trends in popularity of some phenomenon, and will stop observing and commit production capacity to one trend or the other before a production deadline.

Table 1: **Notation Summary**

r	Production rate
t	Time remaining before deadline
v	Percentage of population that seeks vaccination
m	Size of the population
x_t	Observation about X revealed at time t
y_t	Observation about Y revealed at time t
$\theta_{X,t}$	A vector of summary statistics for the trend on X at time t
$\theta_{Y,t}$	A vector of summary statistics for the trend on Y at time t
$\mathcal{X}$	The forecast function for the number of cases of strain X for coming season
$\mathcal{Y}$	The forecast function for the number of cases of strain Y for coming season
$\mathcal{X}_t$	Shorthand notation for $\mathcal{X}(\theta_{X,t})$
$\mathcal{Y}_t$	Shorthand notation for $\mathcal{Y}(\theta_{Y,t})$
$V_t(\theta_{X,t}, \theta_{Y,t})$	Optimal value with t time until deadline

4 Model

In this section, we develop a model that captures the repeated nature of the decision over time: either commit to a course of action or gather new information. We consider a discrete time model, in which the decision can be considered at the beginning of every period. The problem we study is naturally a finite horizon problem because of the deadline.

In each period, the decision maker has a choice between two options we call X and Y. In the vaccine context, X and Y are two candidate strains of the virus for one of the categories. This model captures the possibility of deferring the decision to be considered in the next period while new information becomes available. The fundamental tension is between the benefit of getting more information, to be more certain of selecting the predominant trend, and the sacrifice of reduced production time. We measure value in terms of the expected number of cases of flu prevented.[2]

Table 1 summarizes the notation for the model. Note that we track t, time remaining: the forecast $\mathcal{X}(\theta_{X,t})$ contains one additional period's worth of information compared to the earlier forecast $\mathcal{X}(\theta_{X,t+1})$. See the timeline in Figure 1. Also note that in the vaccine example, the $\mathcal{X}(\theta_{X,t})$ and $\mathcal{Y}(\theta_{Y,t})$ introduced above, the "forecasted" values, are forecasts of cases in the absence

[2]In the vaccine context, there are many aspects to measuring value that we do not capture, including the number of flu deaths prevented by the vaccine, the total cost of the program, as well as the potential side effects from administering the vaccine.

of a vaccine.

Figure 1: Timeline. (FIGURES ATTACHED AT THE END)

The state variables $\theta_{X,t}$ and $\theta_{Y,t}$ are vectors; each contains the summary statistics of the historical data for the respective trend until time t. The particulars of these summary statistics will depend on the specific data collection and belief updating process. At one extreme, the state variables can be the historical data (the set of observations) itself. At the other extreme, the state variables could just be the most up-to-date forecast, a scalar. Between these extremes, we could tally summary statistics such as the current (time t) estimates of parameters of a regression model or the variance of the next prediction. Some, but not necessarily all, of the elements of $\theta_{X,t}$ are used in the forecast for X, $\mathcal{X}(\theta_{X,t})$. 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:

$$V_t(\theta_{X,t}, \theta_{Y,t}) = \max \begin{cases} \frac{\mathcal{X}(\theta_{X,t})}{m} \min\{rt, vm\} & \text{Commit to X now} \\ \frac{\mathcal{Y}(\theta_{Y,t})}{m} \min\{rt, vm\} & \text{Commit to Y now} \\ E\left[V_{t-1}(\tilde{\theta}_{X,t-1}, \tilde{\theta}_{Y,t-1}) | \theta_{X,t}, \theta_{Y,t}\right] & \text{Wait} \end{cases} \tag{1}$$

where the expected value in the wait expression is taken over the future summary statistics $\tilde{\theta}_{X,t-1}$ and $\tilde{\theta}_{Y,t-1}$. The expressions for the "commit" actions assume that the doses are distributed uniformly over the population. If production is a binding constraint: $rt < vm$, then out of the rt units produced, $\frac{\mathcal{X}_t}{m}$ of them will be cases of flu prevented. We are assuming perfect effectiveness of a vaccine dose administered and no cross-effectiveness.

Finally, we define the boundary condition: with no time left, nothing can be produced, resulting in zero value.

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

The dynamic programming equation for the optimal value is the maximum of the values of the three possible actions—commit to one action or the other or wait 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 for the season ($\mathcal{X}$ and $\mathcal{Y}$); instead, we make structural assumptions about the forecast function. (The assumptions for the X trend are shown below; analogous statement for Y

are implied.)

First, we assume that on average, there are not anticipated changes in the forecast. Certainly, there will be changes in the forecast contingent on new information, but the mean value of the next forecast is the current forecast. This assumption implies that all available information is incorporated into the current forecast.

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

A natural conclusion from Assumption 1 is that a higher forecast in one period implies a higher mean in the next period. 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,[3] but rather that good news or bad news now tends to persist in the next period. First-order stochastic dominance can be expressed as a statement about the expected value of increasing functions; we use that convention in stating Assumption 2:

Assumption 2 *For increasing* f, $E[f(\mathcal{X}(\tilde{\theta}_{X,t-1}))|\theta_{X,t}, \mathcal{X}(\theta_{X,t})]$ *is increasing in* $\mathcal{X}(\theta_{X,t})$.

5 Analysis

In this section, we present structural properties of this dynamic decision problem. We start with the structure of the optimal solution to (1) and examine how changes in the production rates and the noisiness of the data affect the decisions.

5.1 Threshold policies

Equation (1) captures the dynamic relationship of the choices over time. The form of the optimal solution can be derived by identifying regions in $(\mathcal{X}_t, \mathcal{Y}_t)$ space for which each of the possible actions is optimal. In other words, finding the optimal solution means that at time t, we say which action is optimal: Commit to X, commit to Y, or wait another period and reconsider. If the vaccine could be produced fast enough so that the total production rt exceeds the number of people who

[3]Random walks and mean-reverting processes both satisfy this property, and neither of those is necessarily increasing.

were interested in getting the vaccine vm, then it would make sense to wait until at least $t = vm/r$ (which would be small if r were large). If, however, $rt < vm$, then further deferral will exacerbate the shortfall in supply.

Figure 2 shows a two-dimensional representation of the optimal solution for this $rt < vm$ case. The solution has a threshold structure. For high enough $\mathcal{X}_t$, commit to X: the cutoff for "high enough" is an increasing function of $\mathcal{Y}_t$. For high enough $\mathcal{Y}_t$, commit to Y: the cutoff for "high enough" is an increasing function of $\mathcal{X}_t$. This is stated more formally in the following proposition:

Figure 2: The form of the optimal solution. (FIGURES ATTACHED AT THE END)

Proposition 1 *For $\mathcal{X}_t \geq \xi_{x,t}(\mathcal{Y}_t)$, commit to X. For $\mathcal{Y}_t \geq \xi_{y,t}(\mathcal{X}_t)$, commit to Y. The thresholds $\xi_{x,t}$ and $\xi_{y,t}$ are increasing. For $\mathcal{X}_t < \xi_{x,t}(\mathcal{Y}_t)$ and $\mathcal{Y}_t < \xi_{y,t}(\mathcal{X}_t)$, wait.*

Proof See Appendix.

The solution to this watchful waiting problem is to defer commitment until there is enough evidence tipping the scales convincingly toward X or Y. The solution strikes a balance between waiting for the evidence to clearly tilt in one direction and leaving time for production. Similar to Wald (1947), McCardle (1985), and Kornish (2004), when the data points definitively one way or the other, commit, otherwise, continue observing.

The concept of enough evidence is related to the endogenous thresholds $\xi_{x,t}$ and $\xi_{y,t}$ and will change as time proceeds. In particular, if the process is stationary (i.e., the observations are generated by a time-invariant process), as the time remaining shrinks, the thresholds become less restrictive for commitment, that is, the continuation region shrinks. In the final period, starting at $t = 1$, the continuation region disappears; because $V_0 \equiv 0$, the optimal action is to commit to X or Y.

Proposition 2 *If the optimal strategy at time t with current forecasts $\mathcal{X}_t$ and $\mathcal{Y}_t$ is to wait, then with the same forecasts at time $t + 1$, (i.e., $\mathcal{X}_{t+1} = \mathcal{X}_t$ and $\mathcal{Y}_{t+1} = \mathcal{Y}_t$) the optimal strategy is to wait.*

Proof See Appendix.

This proposition strikes at the heart of the role of the deadline in this deliberation problem. To understand why the optimal strategy changes as time runs out, we look at the costs and benefits of waiting as time passes. At first, it seems as if the costs of waiting are skyrocketing as the deadline approaches. With one period to go until the deadline, the cost of waiting is "lose everything." That phrasing, while correct, emphasizes the *relative* cost. However, correct comparison is between the absolute benefit and the absolute cost. Unlike the relative cost, the absolute cost of waiting is the same, on average, in every period. The cost is proportional to the lost production (r) in the waiting interval. Therefore, the waiting region shrinks as time passes because of the change in the benefit of waiting; that benefit is declining. The benefit of waiting comes from the possibility that the current best alternatives will be found to be worse and the magnitude of that margin. With a longer set of observations, decreasing variance reduces the expected size of the margin.

Figure 3: The continuation region shrinks as time passes. (FIGURES ATTACHED AT THE END)

5.2 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, due to the egg-derived nature of the product (VRBPAC 2003). One natural question to ask is what is effect of changing the production rate r? Clearly, this will improve the overall value, expected number of cases avoided. But how would such a change affect the decisions?

If r can be increased, but t is small enough so that $rt < vm$, what effect does that have on the optimal strategy? Does this increase make waiting *more* attractive because production is faster in the remaining time? Is it relevant that a higher r raises the stakes in the problem, that is, there are more cases of flu that can be prevented? Does a higher r make waiting *less* attractive because more production is lost (sacrificed) while you wait?

Although each of those questions raises an important issue about changing rates, 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 symmetric (the same for both strains), we find the following

result:

Proposition 3 *For $r < vm/t$, a change in r does not change the optimal strategy.*

Proof See Appendix.

The intuition for this result is as follows. In the $rt < vm$ case, the values of both commit alternatives are proportional to r, and essentially the r "cancels out" in the comparisons.[4] 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).

When the production rate r is the same for X and Y, a higher r does not change their relative attractiveness. In addition, such a change also does not affect the balance between the commit strategies and the wait strategy, because ultimately the wait strategy also rides on the relative attractiveness of the two strains.

But the result above does not hold for the case in which X and Y can be produced at different rates. If instead of a single r, there is a production rate for X (call it r_X) and a production rate for Y (call it r_Y), then we would not expect a null effect from changing one of the rates. Improvement in only one of the rates increases the value of that alternative relative to both of the other alternatives.

Proposition 4 *For $r_X t < vm$ and $r_Y t < vm$, an increase in r_X lowers the threshold $\xi_{x,t}(\mathcal{Y}_t)$ and raises the threshold $\xi_{y,t}(\mathcal{X}_t)$.*

Proof See Appendix.

The proposition can be interpreted to mean that an increase in the rate of production for one of the products has the effect of shifting the continuation region. The product whose rate increases gains territory (i.e., regions of $(\mathcal{X}_t, \mathcal{Y}_t)$), the other product loses territory, and the waiting region can either gain or lose on net. See Figure 4. For some states, the optimal action changes from wait to commit to X; for some the optimal actions changes from commit to Y to commit to X, and for some the optimal action changes from commit to Y to wait. This shift happens because increasing

[4]The objective function is homogeneous in r.

the rate of production of X increases the value of the commit to X strategy the most, then the value of the wait strategy, and does not affect the value of the commit to Y strategy.

Figure 4: Shifts in the optimal regions with an increase in r_X. (FIGURES ATTACHED AT THE END)

In the current environment, the production rate *is* a binding constraint for the flu vaccine decision. With the egg-derived vaccine products, any delay in the vaccine composition recommendation does detract from production in a time period ending at a fixed time. However, one of the key arguments of this analysis is that the loss in production may be worth it to make a more informed choice.

A report from the Institute of Medicine (IOM, 2004) explains how the prominent role of the government as regulator and purchaser of vaccines has made the vaccine supply for all the recommended vaccines quite fragile. There are only three influenza vaccine providers, Aventis Pasteur, Chiron, and MedImmune (VRBPAC 2004, p. 21), which indicates that this market is not a terribly attractive one for the producers.[5] Large firms such as GlaxoSmithKline (GSK) produce the influenza vaccine for many worldwide markets, but not for the U.S. market. The IOM report argues that the legal and regulatory issues stifle innovation and dampen financial incentives in the vaccine markets.

A key takeaway is that substantial increases in the production rate for both products, those that eliminate concerns of shortage, allow a longer period of information gathering. However, smaller increases in the rate (such that production levels remain lower than demand) result in no effect on the deliberation decisions. In contrast, asymmetric changes do have an effect. An increase in the production rate of one of the products will change the optimal strategy at some $(\mathcal{X}_t, \mathcal{Y}_t)$ points (ones that have relatively higher values for the forecast for the product with the increased rate) from wait to buy and other points (ones that have relatively lower values for the forecast for the product with the increased rate) from buy to wait.

[5] And the recent closure of Chiron's plant illustrates this fragility.

5.3 "Noise" in the data stream

The value of the option to defer comes from the fact that new information will be revealed that might change the default choice. In many problems with deferral options, we see that the greater the variability of the underlying stochastic process, the more value there is waiting, as the revealed information has a greater chance of changing the decision. We explore that idea in the context of this model.

For a forecast model with normally distributed errors and a constant (known or unknown) error variance, the error variance nicely represents the variability of the process. To analyze the effect of variance, we break out the state description used in the value function into finer detail. We decompose each state vector ($\theta_{X,t}$ and $\theta_{Y,t}$) into components, the estimates of the regression parameters and the estimate of the regression error variance.

$$\theta_{X,t} \equiv (c_{X,t} \;\; v_{X,t}) \tag{3}$$

$$\theta_{Y,t} \equiv (c_{Y,t} \;\; v_{Y,t}) \tag{4}$$

Note that the c's can be vectors. This separation of the θs into components highlights the two uses of the state information: first to inform the current estimates of the forecasts $\mathcal{X}_t$ and $\mathcal{Y}_t$ and second to create the distribution on future uncertain quantities, used in the evaluation of the wait alternative.

Proposition 5 *If the forecast values $\mathcal{X}(c_{X,t})$ and $\mathcal{Y}(c_{Y,t})$ are convex in the estimates of the forecast parameters, $c_{X,t}$ and $c_{Y,t}$ respectively, and the forecasts are made assuming constant (homoscedastic) normally distributed errors, then increases in the error variances create expansions in the continuation region.*

Proof See Appendix.

One of the conditions of the proposition is that the forecast is convex in the data.[6] This is true in a linear trend model: each observation enters the prediction of a point along the trend line

[6]More generally, the forecast should be convex in the data and the commit strategies should be convex in the forecast. With the value function (1), the commit strategies are linear in the forecast, so that second step is clearly satisfied.

linearly. Given that the season total is the sum of points along the trend line, the linear relationship is preserved. (This convexity condition is not necessarily satisfied by other forecasting forms.)

For a linear trend (see Raiffa and Schlaifer 1961, pp. 57–58), more variability comes from data histories that have lower covariances with t. For $\mathcal{X}$ convex in the data for X and $\mathcal{Y}$ convex in the data for Y, then more uncertainty increases the continuation value (the expected value), which leads to higher thresholds.

5.4 Numerical solution strategy

Like other finite horizon dynamic programming problems, we would like to solve this one by working backward from time 0. However, there are two difficulties that make the "backing up" procedure problematic. The first difficulty is that the expectation expressions cannot be evaluated analytically even for the simplest forecasting form, such as the normal regression process with unknown variance.

The second difficulty is that the state variables are continuous. Therefore, we cannot fully specify the continuation value for all possible values of the state variables. Instead, we need a procedure for estimating the continuation value as a function of the state variables. One way to do this estimation is to use the simulation technique of Longstaff and Schwartz (2001) which they applied to financial options, but can be used for other optimal stopping problems. This technique entails simulating trajectories of the stochastic processes and then using least-squares regressions to estimate the continuation values as a function of the state variables.

6 Recommendations

After studying the influenza composition decision as a quintessential and important example of deliberation with a deadline, we have several recommendations. As was made clear in the transcripts of the VRBPAC deliberations, the option to wait for more information was acknowledged but dismissed. Our first recommendation is to take information collection possibilities more seriously by performing an analysis on a model comparing action with delay such as the one we propose. There are two components that would be essential to following this advice. The first component is the development of prediction models about the spread of the disease. (These are the forecast

functions $\mathcal{X}$ and $\mathcal{Y}$ in our model.[7]) Our understanding is that much of the statistical analysis in this realm is geared toward real-time determination of whether or not the flu season has begun, not on prediction of the size of the epidemic.

One of the objections to work on this type of forecasting is the difficulty of the task; there would be great uncertainty in any predictions. This uncertainty is an essential feature of the problem, and therefore the forecasting mindset must be to emphasize distributions over outcomes and not just the point estimates. It is the uncertainty in the predictions that makes the waiting alternative viable and interesting, so wide ranges of possibility are instructive, not harmful, in this analysis. The second required component is data: to support those 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) sentinel physicians during flu season.

3. Lab samples sent in for virus-typing.

All three of these streams have serious issues with being representative. Would extrapolating to the population level be appropriate from the first stream of data 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 (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?

In addition to the model and data recommendation, we offer two other recommendations, not directly related to the modeling effort in this paper, but important issues that have emerged from the study of this problem. The second recommendation is to pursue increases in production rates

[7] If these models exist, they were not presented in the deliberations.

r. Many in the vaccine community would like to see a culture-derived rather than an egg-derived process for flu vaccine production. While we show that some increases in r don't change the waiting decision, it is clear that they do increase the value measure. As mentioned earlier, the findings in the recent IOM report on the market for vaccines emphasizes that the government's large role in the vaccine markets hinders innovation because of the real or perceived interference with profit. In addition, the current season's supply problems due to the closure of the Chiron plant draws attention to the risks of having so few suppliers. The third recommendation is to revisit the either-or choice. In other words, reexamine the assumptions about the importance of trivalency. If two strains appear to both be strong, why not include them both? We say more about this reframing of the problem in section 7.2.1

7 Discussion

In this discussion, we explore additional questions surrounding deliberation with a deadline beyond the main focus of this paper, information collection on the progression of trends. First we discuss some other types of uncertainties that may show up in this type of problem, and then we propose structural variations on this decision problem, that is, changes to the objective function and the structure of the decisions.

7.1 Other types of uncertainties

As we mentioned in the introduction, there are many analogies between biologic and market phenomenon—in particular the idea of contagion. Another similarity is the pair of uncertainties, technical and market uncertainties. This paper has been all about the so-called market uncertainties, looking at information collection about prevalence. However, there are technical uncertainties in this problem too, such as the uncertainty about the production rate. And with an uncertain production rate, one possible outcome would be an inability to produce at all. This production uncertainty is especially relevant to the flu vaccine problem because of recent deliberations. In 2003, there was an emergent strain, but the committee decided not to include it in the vaccine because the pharmaceutical firms could not produce the vaccine for the new strain reliably. In the deliberations, there seemed to be a general consensus that it was unlikely to happen soon enough,

although there was no talk of anyone's assessments of the probability of short term success. Incorporating this technical or production uncertainty would require these assessments. Experts would have to assess the likelihood of production success over time, as opposed to the very data-oriented tracking task we did include in the model.

Our model has assumed a fixed deadline, but of course in many situations, there would also be uncertainty about deadline. While uncertainty about the deadline (e.g., the start of flu season) makes the problem harder, we acknowledge that the deadline is not as "hard" as we have modeled it. Doses produced early in the season can still be used.

7.2 Structural variations

In this section, we discuss two types of structural variations on the model. In section 7.2.1, we look at a change to the structure of the decisions. In section 7.2.2, we cover changes to the value function.

7.2.1 Implications of the trivalent vaccine

One of the first questions we raised in our discussions of this problem was why does the choice have to be structured as "strain X or strain Y"? Could there be a choice to include both X and Y?

The situation we have studied with this model is the strain choice for a single category of the virus. The flu vaccine used in the U.S. and many other parts of the world is actually a trivalent (three-strain) vaccine—one strain in each of three categories labeled A1, A2, and B. In our analysis, we looked at one strain in isolation. The three strains complicate the decision because while one considers waiting for new data to resolve the A1 prediction (for example), production time is lost on the vaccines for the other strains.[8]

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)? The FDA, the licensing body in the U.S. charged with overseeing the pharmaceutical industry and the arbiter of what is safe and effective, has approved several companies to produce the trivalent formulation in this country. Although the exact virus strain can change each year, the firms still have the approval to produce

[8] Although it has not done so in the past, the FDA could decide on two of the three strains, allowing the firms to start production.

with a new strain filling in for the old. A full cycle of new approvals would take too long to make changes that are responsive to the evolution of the virus. However, a vaccine with four strains in it is not approved. If it were, then in some sense, the choice between contenders X and Y would be simpler: if it's too close to call, include them both. The three-strain inflexibility is a conservative stance: that configuration has many years of safety and efficacy evidence.

An interesting question to look at is what is the optimal number of strains as a function of the information available. Adding more strains gives more disease coverage, but reduces the production rate of full doses due to constrained production capacity. In addition, there are safety and efficacy concerns with additional strains. Examining these trade-offs is the essence of our third recommendation from section 6. Extending the thinking on the number of strains, is there some optimal number of strains—either more or less than three?

7.2.2 Related problem

The problem we have studied is one of deliberation with a deadline. In particular, we have focused on learning about a trend. In this section, we discuss a related problem, which has a similar commit-or-wait structure, but with a slightly different objective function.

The similar problem is as follows: a firm is deciding between two actions, for example, a manager of a line of apparel picking a color scheme or an electronics manufacturer picking a telecommunications standard for a product.[9] In making this choice, the firm is tracking market data about the popularity of the different options, to use in a forecasting model of demand for each of the configurations. Just like in the vaccine problem, the firm faces a repeated commit-or-wait problem, with information revealed in any delay period. The problem formulation is very similar to the vaccine problem described in this paper.

However, one crucial difference is that the sales of the product depend on the level of demand *only* if the production constraint is non-binding. You cannot sell more than is demanded or more than you produce: sales will be the minimum of demand and production. This element of the objective function is different from the vaccine problem in which people do not know if they will be one of the cases of the flu avoided. In the vaccine problem, we forecasted the prevalence of each strain in the absence of the vaccine, and value (cases of flu avoided) is proportional to the

[9] A similar problem is discussed in Fisher et al. 1994.

prevalence. The larger the potential epidemic, the more people are helped for any level of units produced.

The trend-tracking with sales model looks similar to (1):

$$V_t(\theta_{X,t}, \theta_{Y,t}) = \max \begin{cases} min(\mathcal{X}_t, rt) & \text{Commit to X now} \\ min(\mathcal{Y}_t, rt) & \text{Commit to Y now} \\ E\left[V_{t-1}(\tilde{\theta}_{X,t-1}, \tilde{\theta}_{Y,t-1}) | \theta_{X,t}, \theta_{Y,t}\right] & \text{Wait} \end{cases} \tag{5}$$

Notice that the commit alternatives are only proportional to the forecasts for the cases in which the production constraint is not binding (i.e., rt is big). If there is no chance of production (rt) exceeding demand for either product, then it doesn't matter which trend you bet on. Both choices will use up all the production.

The distinguishing element between these two problems is the issue of whether the distribution of the production is targeted (as in the sales model, in which people who demand self-select into purchase) or random (in which the production is distributed to both those who will benefit and those who will not).[10] The random distribution model we study is relevant to mass communication and marketing campaign problems, in which neither the firm nor the consumers knows a priori who will be the people who benefit or are affected by the firm's choice.

8 Conclusions

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. One of the fundamental insights is that while waiting surely has a cost–lost production–it also has a benefit–the opportunity to make a more-informed decision. Of course, these costs and benefits must be compared. In the situations in which the benefit of waiting exceeds the cost of the lost production, one is accepting a more serious shortage in exchange for an increased chance of a better match with the units that are produced.

The model we built and analyzed is a "trend-tracking" model, and we examine the question of when it pays to stop tracking the trend and act. With these two features, this policy decision has analogs in other decision-making realms such as production to meet a deadline like a seasonal

[10] This distinction can be thought of as the distinction between identifiable and statistical beneficiaries.

opportunity or internal scheduling deadline, in a problem domain in which a firm is evaluating two contending approaches, standards, or designs. Neustadt and Fineberg (1982) raise the point that in a political (public) issue, optimal waiting or strategic information gathering can look like indecisive leadership, thus posing an extra challenge for those in the public eye who want to deliberate; the challenge is even greater in a setting with a salient deadline.

These commit-or-wait problems have a similar structure to other sequential analysis stopping problems. When there is strong enough evidence in one direction or the other, commit to that direction. Otherwise, in a "middle region" (literally and figuratively), continue to wait and gather information. Waiting is more attractive time until the deadline is longer or when the degree of uncertainty is high. Interestingly, waiting does not become relatively more or less attractive when production rates change symmetrically.

Many, many analyses go in to supporting the decisions and deliberations of the VRBPAC. Data is collected, analyzed, and pored over: 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. Hurdles to applying this technique were discussed in section 6. In particular, 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 don't 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.

We have used this public sector problem to motivate an analysis which is at the heart of some of the most challenging and important issues in new product development efforts. New products are conceived to meet burgeoning needs in a consumer base. Whenever there is something new, there is someone who wants to know if this is an important change for the future, or is this a flash in the pan? Is this one of Christensen's "disruptive technologies" or will it be a short-lived fad? Will this new idea, trend, or product reach the vaunted tipping point (see, e.g., Gladwell 2000, Granovetter and Soong 1983)? There are fundamental questions in many realms, and our aim is to contribute

to the thinking on them with this exploration of a sequential decision problem under uncertainty.

9 Appendices

9.1 Proof of Proposition 1

We start by showing that for a given $\mathcal{Y}_t$, there is a cutoff $\xi_{x,t}(\mathcal{Y}_t)$ such that for $\mathcal{X}_t \geq \xi_{x,t}(\mathcal{Y}_t)$, it is optimal to commit to X now. We do this by showing that above a threshold on the current forecast for X, there is no difference between the value of the optimal strategy and the value of the commit to X strategy. In other words, show that

$$V_t(\theta_{X,t}, \theta_{Y,t}) - \frac{\mathcal{X}_t}{m} rt \tag{6}$$

is decreasing in $\mathcal{X}_t$.

Use induction on the number of periods left.

1. First show it holds for $t = 1$. It does because

$$\max\{\mathcal{X}_1 r/m, \mathcal{Y}_1 r/m, 0\} - \mathcal{X}_1 r/m$$

   is decreasing in $\mathcal{X}_1$.

2. Now assume the claim is true for $t-1$:

$$V_{t-1}(\theta_{X,t-1}, \theta_{Y,t-1}) - \frac{\mathcal{X}_{t-1}}{m} r(t-1) \tag{7}$$

   is decreasing in $\mathcal{X}_{t-1}$.

3. Show that it follows that the claim holds for t.

   Show that

$$\max\{\mathcal{X}_t rt/m, \mathcal{Y}_t rt/m, E[V_{t-1}(\theta_{X,t-1}, \theta_{Y,t-1})|\theta_{X,t}, \theta_{Y,t}]\} - \mathcal{X}_t rt/m \tag{8}$$

   is decreasing in $\mathcal{X}_t$.

If either of the first two argument of the max{} expression yields the maximum, the claim obviously holds. The third argument needs further attention; show that

$$E[V_{t-1}(\theta_{X,t-1}, \theta_{Y,t-1})|\theta_{X,t}, \theta_{Y,t}] - \frac{\mathcal{X}_t}{m} rt \tag{9}$$

is decreasing in $\mathcal{X}_t$. Rewrite that expression by adding and subtracting $E\left[\frac{\mathcal{X}_{t-1}}{m}r(t-1)|\theta_{X,t},\theta_{Y,t}\right]$ and then rearranging to get

$$E\left[V_{t-1}(\theta_{X,t-1},\theta_{Y,t-1})-\frac{\mathcal{X}_{t-1}}{m}r(t-1)|\theta_{X,t},\theta_{Y,t}\right]+E\left[\frac{\mathcal{X}_{t-1}}{m}r(t-1)|\theta_{X,t},\theta_{Y,t}\right]-\frac{\mathcal{X}_t}{m}rt \tag{10}$$

The expression inside the first expectation is decreasing in $\mathcal{X}_{t-1}$ by the induction hypothesis. Using Assumption 2, we conclude that the first term (i.e., the first expectation term) is decreasing in $\mathcal{X}_t$. Now it suffices to show that $E\left[\frac{\mathcal{X}_{t-1}}{m}r(t-1)|\theta_{X,t},\theta_{Y,t}\right]-\frac{\mathcal{X}_t}{m}rt$ is decreasing in $\mathcal{X}_t$. By Assumption 1, the first term is equivalent to $\frac{\mathcal{X}_t}{m}r(t-1)$, so the difference is $-\mathcal{X}_t r/m$, which is decreasing in $\mathcal{X}_t$.

The exact same logic holds for $\mathcal{Y}_t$: for a given $\mathcal{X}_t$, there is a cutoff $\xi_{y,t}(\mathcal{X}_t)$ such that for $\mathcal{Y}_t$ above it, it is optimal to commit to Y now.

Finally, we show that the threshold functions are increasing: $\xi_{x,t}(\mathcal{Y}_t)$ is increasing in $\mathcal{Y}_t$. The threshold $\xi_x(\mathcal{Y}_t)$ is maximum of the two $\mathcal{X}_t$s at the intersections of $\frac{\mathcal{X}_t}{m}rt=\frac{\mathcal{Y}_t}{m}rt$ and $\frac{\mathcal{X}_t}{m}rt=E\left[V_{t-1}(\tilde{\theta}_{X,t-1},\tilde{\theta}_{Y,t-1})|\theta_{X,t},\theta_{Y,t}\right]$ The $\mathcal{X}_t$ solving the first equation is clearly increasing in $\mathcal{Y}_t$. The $\mathcal{X}_t$ solving the second equation is also increasing in $\mathcal{Y}_t$ by Assumption 2 and the fact that V_t is increasing in $\mathcal{X}_t$ and $\mathcal{Y}_t$.

9.2 Proof of Proposition 2

At time t, pick a point in $(\mathcal{X}_t,\mathcal{Y}_t)$ space such that it is optimal to wait. That is,

$$\begin{aligned}\frac{\mathcal{X}_t}{m}rt &\leq E\left[V_{t-1}(\tilde{\theta}_{X,t-1},\tilde{\theta}_{Y,t-1})|\theta_{X,t},\theta_{Y,t}\right], \text{ and}\\ \frac{\mathcal{Y}_t}{m}rt &\leq E\left[V_{t-1}(\tilde{\theta}_{X,t-1},\tilde{\theta}_{Y,t-1})|\theta_{X,t},\theta_{Y,t}\right]\end{aligned}$$

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

$$\frac{\mathcal{X}_{t+1}}{m}r(t+1) \leq E\left[V_t(\tilde{\theta}_{X,t},\tilde{\theta}_{Y,t})|\theta_{X,t+1},\theta_{Y,t+1}\right], \text{ and} \tag{11}$$

$$\frac{\mathcal{Y}_{t+1}}{m}r(t+1) \leq E\left[V_t(\tilde{\theta}_{X,t},\tilde{\theta}_{Y,t})|\theta_{X,t+1},\theta_{Y,t+1}\right] \tag{12}$$

From the condition in the proposition, $\mathcal{X}_{t+1}=\mathcal{X}_t$ and $\mathcal{Y}_{t+1}=\mathcal{Y}_t$, so the LHSs of (11) and (12) are $\frac{\mathcal{X}_t}{m}r(t+1)$ and $\frac{\mathcal{Y}_t}{m}r(t+1)$, respectively. It suffices to show that

$$E\left[V_t(\tilde{\theta}_{X,t},\tilde{\theta}_{Y,t})|\theta_{X,t+1},\theta_{Y,t+1}\right]\geq E\left[V_{t-1}(\tilde{\theta}_{X,t-1},\tilde{\theta}_{Y,t-1})|\theta_{X,t},\theta_{Y,t}\right]+\frac{\mathcal{X}_t}{m}r \tag{13}$$

(and similarly for Y).

This proposition examines a situation for which when there are t periods remaining, waiting is better than committing. Now add an additional period on to the end of the horizon: this gives us a comparison of starting from the same initial state $(\mathcal{X}_t, \mathcal{Y}_t)$, but with $t+1$ periods to go instead of t. With the extra period, the commit alternatives improve in a straightforward manner. The commit to X alternative improves by one additional period's worth of production: $\mathcal{X}_t r/m$, and the commit to Y alternative improves by $\mathcal{Y}_t r/m$.

Now we argue that the wait alternative improves by at least as much, on average, as either one of those amounts. After waiting in the initial period, if the decision maker follows the t-period horizon strategy (which is possibly but not necessarily optimal for the $t+1$-period horizon), then with the additional period at the end, one period's worth of extra value will be realized either for production for X or Y. By Assumption 1, the expected values of any future period's forecasts are equal to the current values. Considering the waiting alternative: if the production in the final period ends up being X, the expected value of that additional production is $\mathcal{X}_t r/m$, and if the production in the final period ends up being Y, the expected value of that additional production is $\mathcal{Y}_t r/m$. In addition, under the wait in the initial period alternative, the value from the extra period is increased by the flexibility to choose the higher valued strategy—so the expected value of additional production in the extra period is greater than $\mathcal{X}_t r/m$ and $\mathcal{Y}_t r/m$.

Therefore, because it was optimal to wait with t periods remaining, and with the addition of an extra period, the value of the wait alternative increases even more than the value of the commit alternatives, it is optimal to wait with $t+1$ periods remaining.

9.3 Proof of Proposition 3

Proof that changes in r do not affect the optimal buy vs. wait decisions for $rt < vm$. The proof is by induction, showing that at time t, the strategies are invariant to a change in r.

1. Show the result holds with one period to go, $t = 1$: $V_1(\theta_{X,1}, \theta_{Y,1}) = \max\left\{\frac{\mathcal{X}_1}{m} r1, \frac{\mathcal{Y}_1}{m} r1, 0\right\} = r \max\left\{\frac{\mathcal{X}_1}{m}, \frac{\mathcal{Y}_1}{m}, 0\right\}$. The comparison between the three alternatives does not change as r changes; e.g., the set of $(\mathcal{X}_1, \mathcal{Y}_1)$ for which $\frac{\mathcal{X}_1}{m} r1 \geq \frac{\mathcal{Y}_1}{m} r1$ is not affected by r.

2. Assume the result holds with $t-1$ periods remaining. In other words, assume the three comparisons between the alternatives do not depend on r with $t-1$ periods remaining because

the r can be factored out. Define

$$W_{t-1}(\theta_{X,t-1},\theta_{Y,t-1}) = \max \begin{cases} \frac{\mathcal{X}_{t-1}}{m}(t-1) \\ \frac{\mathcal{Y}_{t-1}}{m}(t-1) \\ E\left[W_{t-2}(\tilde{\theta}_{X,t-2},\tilde{\theta}_{Y,t-2})|\theta_{X,t-1},\theta_{Y,t-1}\right] \end{cases} \tag{14}$$

and assume $V_{t-1} = rW_{t-1}$.

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

$$W_t(\theta_{X,t},\theta_{Y,t}) = \max \begin{cases} \frac{\mathcal{X}_t}{m}(t) \\ \frac{\mathcal{Y}_t}{m}(t) \\ E\left[W_{t-1}(\tilde{\theta}_{X,t-1},\tilde{\theta}_{Y,t-1})|\theta_{X,t},\theta_{Y,t}\right] \end{cases} \tag{15}$$

$$rW_t(\theta_{X,t},\theta_{Y,t}) = \max \begin{cases} \frac{\mathcal{X}_t}{m}(rt) \\ \frac{\mathcal{Y}_t}{m}(rt) \\ rE\left[W_{t-1}(\tilde{\theta}_{X,t-1},\tilde{\theta}_{Y,t-1})|\theta_{X,t},\theta_{Y,t}\right] \end{cases} \tag{16}$$

Because r is a constant, it can be moved inside the expectation operator in the third term of equation (16), so that term becomes $E\left[rW_{t-1}(\tilde{\theta}_{X,t-1},\tilde{\theta}_{Y,t-1})|\theta_{X,t},\theta_{Y,t}\right]$. Buy the induction hypothesis, $V_{t-1} = rW_{t-1}$, so equation (16) can be rewritten as

$$rW_t(\theta_{X,t},\theta_{Y,t}) = \max \begin{cases} \frac{\mathcal{X}_t}{m}(rt) \\ \frac{\mathcal{Y}_t}{m}(rt) \\ E\left[V_{t-1}(\tilde{\theta}_{X,t-1},\tilde{\theta}_{Y,t-1})|\theta_{X,t},\theta_{Y,t}\right] \end{cases} \tag{17}$$

which is equal to V_t.

9.4 Proof of Proposition 4

Proof that an increase in the production rate for X, r_X, lowers the threshold on $\mathcal{X}_t$ above which commit to X is optimal and raises the threshold on $\mathcal{Y}_t$ above which commit to Y is optimal.

With different production rates, for the case $r_X t < vm$ and $r_Y t < vm$, the optimal value function (1) becomes

$$V_t(\theta_{X,t},\theta_{Y,t}) = \max \begin{cases} \frac{\mathcal{X}_t}{m} r_X t & \text{Commit to X now} \\ \frac{\mathcal{Y}_t}{m} r_Y t & \text{Commit to Y now} \\ E\left[V_{t-1}(\tilde{\theta}_{X,t-1},\tilde{\theta}_{Y,t-1})|\theta_{X,t},\theta_{Y,t}\right] & \text{Wait} \end{cases} \tag{18}$$

The threshold on $\mathcal{X}_t$, ξ_x, above which it is optimal to commit to X, is determined by two comparisons:

$$\frac{\mathcal{X}_t}{m} r_X t \;\; > \;\; \frac{\mathcal{Y}_t}{m} r_Y t \tag{19}$$

$$\frac{\mathcal{X}_t}{m} r_X t \quad > \quad E\left[V_{t-1}(\tilde{\theta}_{X,t-1}, \tilde{\theta}_{Y,t-1}) | \theta_{X,t}, \theta_{Y,t}\right] \tag{20}$$

In the first comparison (19), the LHS is increasing in $\mathcal{X}_t$ and r_X and the RHS is not a function of either. See Figure 5. Therefore, an increase in r_X shifts the LHS up, reducing the $\mathcal{X}_t$ at which the LHS and RHS intersect.

Figure 5: Figures used in the Proof of Proposition 4. (FIGURES ATTACHED AT THE END)

Figure 6: Figures used in the Proof of Proposition 4. (FIGURES ATTACHED AT THE END)

In the second comparison (20), both the LHS and RHS are increasing in $\mathcal{X}_t$ and r_X. The LHS has a slope of $r_X t/m$ with respect to $\mathcal{X}_t$. Because the RHS, the value of waiting, is a convex combination of future commit values, we know that the slope of the RHS with respect to $\mathcal{X}_t$ cannot be more than $r_X(t-1)/m$. Therefore, the slope of the LHS is greater than the slope of the RHS w.r.t. $\mathcal{X}_t$. As r_X increases, both slopes increase. The slope of the LHS changes by t/m for a unit increase in r_X. The slope of the RHS changes by at most $(t-1)/m$. Referring to Figure 5, we see that the point of intersection of the two curves must decrease.

The threshold for $\mathcal{X}_t$, above which it is optimal to commit to X is determined by the $\mathcal{X}_t$ that satisfy both (19) and (20). The threshold $\xi_{x,t}$ is the maximum of the cutoffs for those two conditions. If both cutoffs are lowered, then the maximum of the two cutoffs is also lowered. An increase in r_X lowers the cutoffs for both conditions, so the threshold $\xi_{x,t}$ is reduced with an increase in r_X.

The argument for the effects on the threshold on $\mathcal{Y}_t$, $\xi_{y,t}$, is analogous. There are two comparisons. The first comparison is the same as (19) with the inequality reversed. The cutoff on $\mathcal{Y}_t$ increases because it is determined by the intersection of a function increasing in $\mathcal{Y}_t$ (i.e., $\frac{\mathcal{Y}_t}{m} r_Y t$) and a constant in $\mathcal{Y}$ (i.e., $\frac{\mathcal{X}_t}{m} r_X t$). See Figure 5. An increase in r_X increases the constant, so the point of intersection increases. The second comparison is as follows.

$$\frac{\mathcal{Y}_t}{m} r_Y t \quad > \quad E\left[V_{t-1}(\tilde{\theta}_{X,t-1}, \tilde{\theta}_{Y,t-1}) | \theta_{X,t}, \theta_{Y,t}\right] \tag{21}$$

Here, the LHS is increasing in $\mathcal{Y}_t$ but not a function of r_X. The RHS is also increasing in $\mathcal{Y}_t$, but is an increasing function of r_X. So an increase in r_X raises the slope of the RHS, increasing the point of intersection.

9.5 Proof of Proposition 5

Rewriting (1) using the decomposition from (3) and (4),

$$V_t(\mathcal{X}(c_{X,t}), \mathcal{Y}(c_{X,t}), v_{X,t}, v_{Y,t}) = \max \begin{cases} \frac{\mathcal{X}_t}{m} \min\{rt, vm\} & \text{Commit to X now} \\ \frac{\mathcal{Y}_t}{m} \min\{rt, vm\} & \text{Commit to Y now} \\ E\left[V_{t-1}(\tilde{\mathcal{X}}_{t-1}, \tilde{\mathcal{Y}}_{t-1}, \tilde{v}_{X,t-1}, \tilde{v}_{Y,t-1}) | \mathcal{X}_t, \mathcal{Y}_t, v_{X,t}, v_{Y,t}\right] & \text{Wait} \end{cases} \tag{22}$$

First show that V_t is convex in the forecast estimate $\mathcal{X}_t$ (and the same argument holds for $\mathcal{Y}_t$) by induction. In period 1, $V_1 = \max\{\mathcal{X}_1 r/m, \mathcal{Y}_1 r/m, 0\}$ which is convex in $\mathcal{X}_1$.

Now assume V_{t-1} is convex in $\mathcal{X}_{t-1}$. Show it follows that V_t is convex in $\mathcal{X}_t$. The expression for the value of the "wait" alternative is the expected value of a function we have assumed is convex (by the inductive assumption). The expected value of a convex function is also convex because it is a positive-weighted sum of convex functions. So the "wait" alternative is convex in $\mathcal{X}_t$. Further, the value of the "commit to X" alternative is linear (and therefore convex) in $\mathcal{X}_t$, and the value of the "commit to Y" alternative is not a function of $\mathcal{X}_t$. The maximum of convex functions is convex, so by induction, V_t is convex in $\mathcal{X}_t$.

Second show that the value of the wait alternative is increasing in the variance of the forecast. We will combine that claim with the fact that the value of the other two alternatives are not functions of the variance of the forecast to conclude that the continuation region expands with an increase in the variance.

To show that the value of the wait alternative is increasing in the variance of the forecast, we use the result that a mean-preserving spread increases the expected value of a convex function of a random variable (see, e.g., Rothschild and Stiglitz, 1970). Given that the forecasts are made using normal errors, an increase in the variance of the error is a mean-preserving spread. We showed above that V_t is convex in $\mathcal{X}_t$. A higher variance of the regression error makes increases the variance of the prediction of any point on the trend, and therefore increases the variance of $\mathcal{X}_t$, comprised as the sum of points along the trend.

Increasing the value of the waiting strategy without changing the values of the other strategies expands the region over which waiting is optimal.

References

[1] Frank M. Bass. A new product growth model for consumer durables. *Management Science*, 30(January):215–227, 1969.

[2] Clayton L. Christensen. *The Innovator's Dilemma*. Harvard Business School Press, Boston, Massachusetts, 1997.

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

[4] CNN.com. WHO plans flu pandemic summit. *CNN*, November 1, 2004. URL: http://www.cnn.com/2004/HEALTH/10/31/flu.pandemic.ap/index.html.

[5] Marshall L. Fisher, Janice H. Hammond, Walter R. Obermeyer, and Ananth Raman. Making supply meet demand in an uncertain world. *Harvard Business Review*, (May-June):83–93, 1994.

[6] M.A. Girshick. Contributions to the theory of sequential analysis. *The Annals of Mathematical Statistics*, 17, 1946.

[7] Malcolm Gladwell. *The Tipping Point*. Little, Brown and Company, Boston, Massachusetts, 2000.

[8] Mark Granovetter and R. Soong. Threshold models of diffusion and collective behavior. *Journal of Mathematical Sociology*, 9:165–179, 1983.

[9] Richard Jensen. Adoption and diffusion of an innovation of uncertain profitability. *Journal of Economic Theory*, 27(1):182–193, 1982.

[10] Laura J. Kornish. Technology choice and timing with positive network effects. Working paper, Duke University, 2004.

[11] Steven A. Lippman and John J. McCall. The economics of job search: A survey. *Economic Inquiry*, 14(2):155–189, 1976.

[12] Francis A. Longstaff and Eduardo S. Schwartz. Valing american options by simulation: A simple least-squares approach. *The Review of Financial Studies*, 14(1):113–147, 2001.

[13] Kevin F. McCardle. Information acquisition and the adoption of new technology. *Management Science*, 31(11):1372–1389, 1985.

[14] Richard E. Neustadt and Harvey Fineberg. *The Epidemic that Never Was: Policy-Making and the Swine Flu Scare*. Random House, Inc., New York, 1982.

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

[16] Howard Raiffa and Robert Schlaifer. *Applied Statistical Decision Theory*. The M.I.T. Press, Cambridge, Massachusetts, 1961.

[17] Everett M. Rogers. *Diffusion of Innovations*. The Free Press, Ontario, Canada, 1962.

[18] Everett M. Rogers. *Diffusion of Innovations, Fourth Edition*. The Free Press, New York, 1995.

[19] Michael Rothschild and Joseph E. Stiglitz. Increasing risk I: A definition. *Journal of Economic Theory*, pages 225–243, 1970.

[20] John Treanor. Weathering the influenza vaccine crisis. *New England Journal of Medicine*, 351(20):2037–2040, 2004.

[21] VRBPAC. Transcript of Vaccines and Related Biological Products Advisory Committee Meeting. URL: http://www.fda.gov/ohrms/dockets/ac/04/transcripts/4020t1.htm, February 18, 2004.

[22] VRBPAC. Transcript of Vaccines and Related Biological Products Advisory Committee Meeting. URL: http://www.fda.gov/ohrms/dockets/ac/03/transcripts/3941t1.doc, March 18, 2003.

[23] Abraham Wald. *Sequential Analysis*. Wiley Publications in Statistics, New York, 1947.

[24] Jeanne Whalen, Betsy McKay, and Sarah Lueck. U.S. flu vaccines face rationing as plant is shut. *Wall Street Journal*, page A1, October 6, 2004.

[25] Joseph T. Wu, Lawrence M. Wein, and Alan S. Perelson. Optimization of influenza vaccine selection. *Operations Research*, to appear.

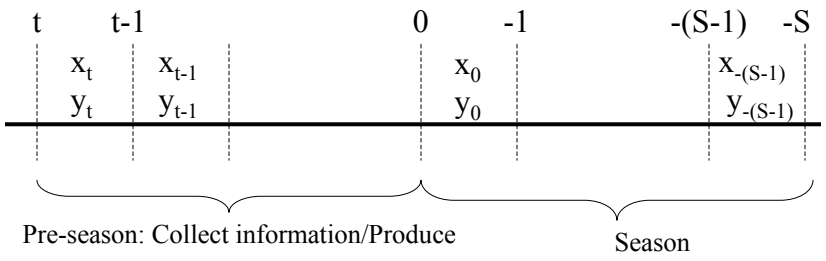


Figure 1: Timeline

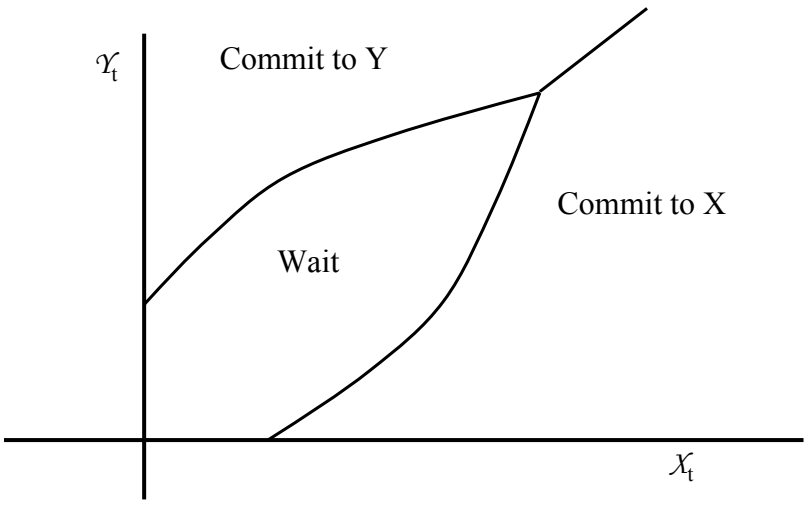


Figure 2: The form of the optimal solution.

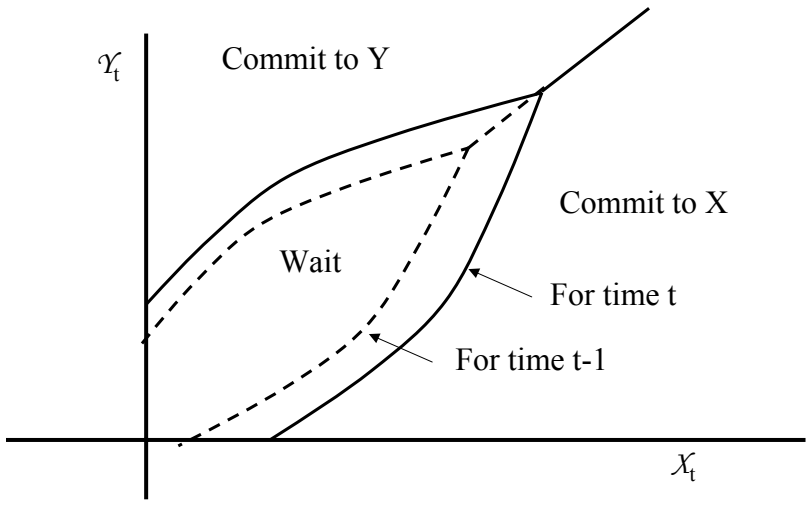


Figure 3: The continuation region shrinks as time passes.

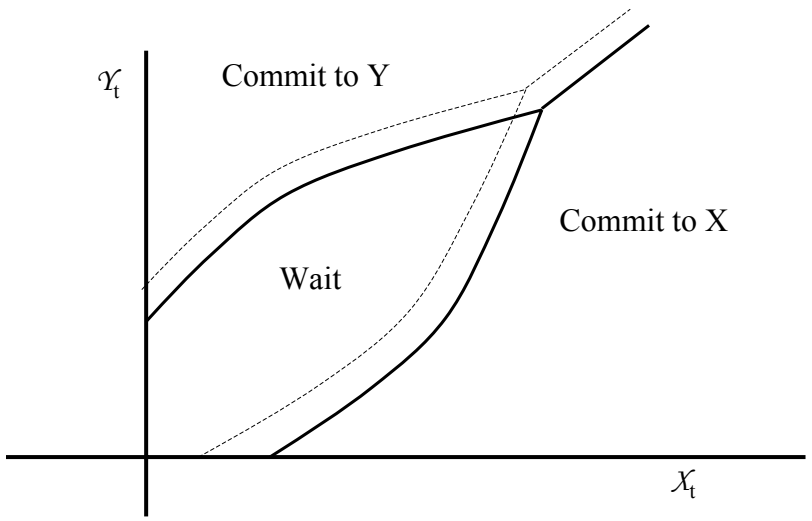
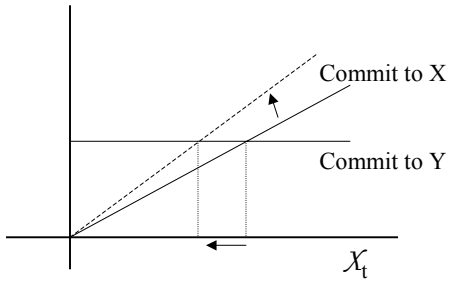
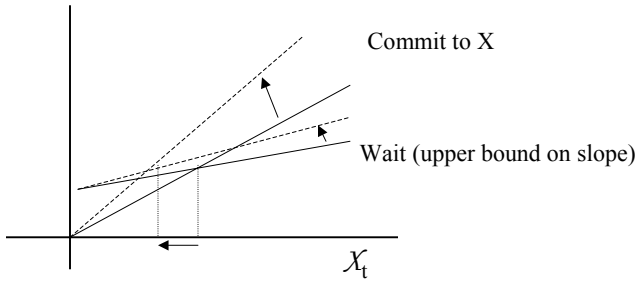


Figure 4: Shifts in the optimal regions with an increase in r_X.

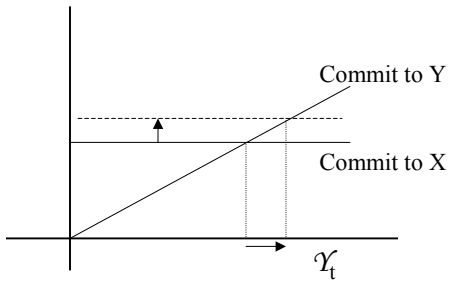


An increase in r_X lowers the cutoff for the comparison between commit to X and commit to Y.

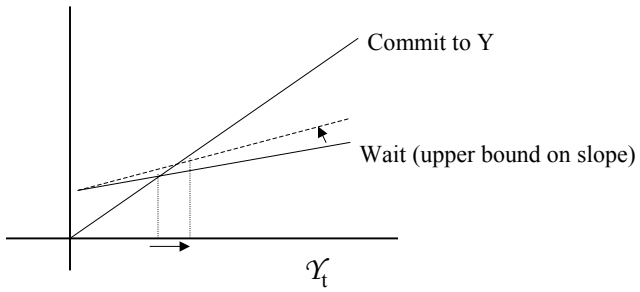


An increase in r_X lowers the cutoff for the comparison between commit to X and wait.

Figure 5: for proof of effects of asymmetric changes in production rates.



An increase in r_X raises the cutoff for the comparison between commit to X and commit to Y.



An increase in r_X raises the cutoff for the comparison between commit to Y and wait.

Figure 6: for proof of effects of asymmetric changes in production rates.