INSEAD
R & D

DECISIONS UNDER PRELIMINARY INFORMATION:
RUSH AND BE WRONG OR WAIT AND BE LATE?

by

C. LOCH*
and
C. TERWIESCH**

2000/83/TM

* Associate Professor of Operations Management at INSEAD, Boulevard de Constance, 77305 Fontainebleau Cedex, France.

** The Wharton School, University of Pennsylvania, Philadelphia, USA.

Decisions Under Preliminary Information: Rush and Be Wrong or Wait and Be Late?

Christoph H. Loch
INSEAD

Christian Terwiesch
The Wharton School

November 2000

Abstract

Many decisions in Operations management involve a choice between early action at low cost but with little information and delayed action with full information, but at a higher cost. For example, in managing a supply-chain, one can postpone some orders until more demand information becomes available. This decision, also known as Quick Response, has been analyzed as a multi-stage newsvendor model with Bayesian updating. A similar problem exists in the context of product development, where often design decisions have to be made despite residual uncertainty about market requirements or inputs from other information providing activities. In this context, however, the uncertain information is not a continuous parameter (such as demand), but a discrete set of non-comparable (design) alternatives. Thus, existing models from the supply-chain literature are not applicable. In this article, we develop a generalized model of decision making based on preliminary information and present applications in the fields of product design and supply chain management.

1 Introduction

Many decisions in product development and supply chain management require a choice between an early action at low cost but with little information (and thus the danger of error), and a delayed action with full information but at substantially higher cost. If the information is made available gradually, rather than in one single step, the decision maker can act at several points in time.

She can either act before receiving any information, relying on *prior information*, which leads to low cost for any given action, but a high risk of a mismatch between what is done and what, *ex-post*, should have been done: parts are ordered for a configuration that the customer ultimately decides not to order, tools are developed for drawings that are later modified by a design department, or orders are placed for ski-parkas that are not matching the consumer's taste.

In the other extreme, the decision maker can choose to delay any action until *all* uncertainty is resolved. While this avoids errors and mismatch, it may make a response prohibitively expensive: activities delay the critical path, and orders may need expediting or slip beyond the quoted lead-time.

As an intermediate option, the decision maker can act after receiving *preliminary information*, which allows a more accurate although not perfect response. Such preliminary information can be given in the form of early-season sales in managing an apparel supply chain, of preliminary drawing releases in concurrent development teams, or of initial discussions with a customer in make-to-specification production.

The sharing of preliminary information is facilitated by recent advances in communication technologies and by closer links within and across organizations. This is true in supply chains, where preliminary information about forecasts is shared through vertically integrated information systems, as well as in product development, where easily transferrable digital design representations have almost completely replaced the traditional libraries of paper drawings. However, even in a world of almost unlimited communication bandwidth, there remains a fundamental problem: the more visibility one party gains into the ongoing information processing activities of another party, the more likely the resulting information is preliminary rather than final. Thus, making decisions based on such preliminary information becomes a key managerial problem.

Decisions based on preliminary information and Bayesian updating have received much attention in the supply chain literature (e.g., Fisher & Anand 1996, Donohue 1996, Eppen & Iyer 1997). However, other environments, in particular product development, have not yet been analyzed with a comparable methodology. The major limitation of applying the existing models to these settings lies in a fundamental difference in the topology of the underlying decision and outcome spaces. How much to order (decision) and how much demand is later realized (outcome) can all be described in a one-dimensional, ordered space, including a distance measure between and among actions and outcomes. Other environments where preliminary information is exchanged are often more complex.

Consider the now common situation of a product development team that applies Concurrent Engineering, where multiple interdependent activities are performed in parallel to shorten project lead time (e.g., Krishnan *et al.* 1997, Loch & Terwiesch 1998). Thus, one subteam begins its work based on preliminary information from another subteam. The decision of how much engineering effort to invest based on preliminary information has similar features to forecast-sharing in a supply chain. However, this preliminary information rarely exhibits the one-dimensional structure of existing Bayesian updating models.

This paper proposes a unifying framework of making decisions under preliminary information and updating. Our analysis yields three important insights. **First**, building on work by Marschak & Radner (1972), we relax the assumption of a single, ordered dimension of actions and outcomes to the more general topology of a sigma algebra that is refined over time. This allows for a whole array of new managerial applications, two of which we discuss explicitly: product design and make-to-specification production.

Second, we show for a general probability space how the optimal response to preliminary information depends on the amount of information available as well as on the cost functions.

Third, the structure of the underlying problem influences to what extent one should commit to an action under preliminary information. In addition to the one-dimensional, ordered structure of the newsvendor problem, we identify three other problem structures frequently encountered in operations management: unstructured, decomposed and hierarchical. We also demonstrate how the traditional one-dimensional structure can be viewed as a special case of our framework. Each problem structure can be modeled as a two-stage stochastic dynamic programming problem. We solve each of these problems in closed form, leading to a number of generic strategies of responding to preliminary information (hedging, pipelining, postponement, templating).

The paper is organized as follows. We first review literature related to preliminary information, emphasizing the congruence between situations in supply chain management, product development, and build-to-specification. Section 3 presents the general problem formulation and our main theoretical result. Section 4 discusses specific problem structures and identifies managerial strategies of responding to preliminary information.

2 Literature

Figure 1: Making decisions under preliminary information

Figure 1 illustrates the generic structure of decision-making under preliminary information. Based on prior information, some actions can be carried out prior to receiving preliminary information. At this point, costs are relatively low. Then the decision maker receives preliminary information in the form of a signal about the final outcome. Using Bayesian updating, she can revisit her initial decision set, possibly leading to additional action at higher cost. Finally, the uncertainty is resolved and may reveal a mismatch between actions and outcome. Corrective action at this point is the most expensive.

The Quick Response (QR) movement in the supply chain literature (Hammond 1990) provides an example of this problem structure. A buyer (e.g. a fashion retailer) places purchase orders over certain quantities to a manufacturer, prior to knowing market demand. The uncertainty inherent in this decision forces the buyer to trade off the cost of ordering too much (leading to excess inventories and mark-downs) with the cost of ordering too little (lost opportunities of making a profit). This decision can be modeled by the classical "newsvendor" model (e.g., Nahmias 1993), extended to a series of information exchanges, with repeated updating and refinement until the final purchase commitment.

Such a sequence of information exchanges allows the supply chain to delay the time of the final quantity commitment until additional demand information becomes available. This is typically modeled as Bayesian updating (e.g., Cachon 1999). If the manufacturer is sure of some portion of the overall demand and can, thus, produce it before the revised forecast, production and shipment costs can be reduced. QR models with updating of quantity forecasts have led to a number of powerful insights (e.g., Fisher and Raman 1996, Donohue 1996, Eppen and Iyer 1997). These models rest on an important assumption: the underlying decision space is one-dimensional, ordered, and that there exists a distance measure between and among actions (how much to order) and outcomes (demand).

The concept of preliminary information is applicable to contexts other than supply chains.

Influential work by McCardle (1985) and McCardle & Lippman (1987) applied information economics to R&D decisions with Bayesian updating of information about future profits. They showed in a simple model that a better preliminary information (a sharper signal) may *delay* the decision to adopt an innovation.

In the context of product development, Clark and Fujimoto (1989) demonstrated empirically that an early start of an information-receiving development activity reduces project lead-time. This work has initiated widespread interest in the management of preliminary information within the field of product development. For example, Krishnan *et al.* (1997) operationalized preliminary information as the possible range of a design parameter (e.g., the depth of a door handle). The speed at which the parameter range can be narrowed, together with the cost of responding to modifications, influences how much design work should proceed in parallel. Loch & Terwiesch (1998) modeled preliminary information as a stream of engineering changes, explicitly including the increasing cost of action over time. They derived the optimal starting point at which downstream should commence its work.

The fundamental difference between the information exchanged in product development and in a supply chain lies in the topology of the underlying decision and outcome spaces. Terwiesch *et al.* (1999) describe five cases of preliminary information sharing in an automotive development project. In one case, the preliminary information is communicated in form of an unordered set with just a handful of outcomes (e.g., different configurations of an automotive climate control system). In other cases, the information resembles more closely the postponement situation described by Lee & Tang (1997): some actions are common over all scenarios and can thereby be executed early (at low cost), while others require a complete resolution of uncertainty and are better delayed as much as possible.

Similarly, Cohen *et al.* (2000) describe the case of a semiconductor equipment manufacturer which builds equipment to order for a semiconductor manufacturer. While the interactions between the two parties again are similar to the QR situation, the preliminary information exchanged is less about the (one-dimensional) order quantity, but more about the (multi-dimensional) specification and configuration of the equipment. A large variety of product configurations, each of which is using different sub-assemblies, forces the manufacturer to choose which (long lead-time) components to order from the 2nd tier supplier. In many cases, the equipment manufacturer needs to make these sourcing decisions prior to having a firm purchase order or detailed specifications. This requires trading off the cost of ordering sub-assemblies too late (after the firm purchase order), in which case the equipment lead-time lengthens, with the cost of ordering them too early, in which case some sub-assemblies might be ordered but not used.

3 A General Model of Preliminary Information

We start the definition of our model with a probability space $(\Omega, \mathcal{F}, P)$ summarizing all possible events, or outcomes of the uncertainty, and the initial knowledge about the outcome (such as the final demand for the newsvendor, the ultimately desired machine configuration in build-to-specs, or the final outcome of an upstream design activity). To keep the structure of the model simple, we assume that $\mathcal{F}$ is countable. This is sufficient to include both the newsvendor application and the product development applications, where significant design alternatives are often discrete and finite in number.

Let $\{A_i; i = 1, ...\}$ be the set of events generating $\mathcal{F}$. We call these "basic"; they cannot be further subdivided into distinguishable events. Analogous to Marschak & Radner (1972, 53 - 59), let $\Psi \subset \mathcal{F}$ be a "coarser" sigma field containing preliminary information about the uncertain outcome. Within Ψ, a signal is conveyed after period one. Let $\{B_i; i = 1, ...\}$ be the set of events generating Ψ. Unions of these basic events are the possible signals that can be perceived as preliminary information. Ψ being coarser than $\mathcal{F}$ implies that any B_j must be the union of (one or) several A_i.

Figure 2: The general model

The sequence of actions is shown in detail in Figure 2. An action $a_1 \in \mathcal{F}$ "addresses" a certain set of possible events. In the example of Cohen et al. described above, an action could describe one or more machine configurations that the equipment manufacturer decides to prepare for by ordering the corresponding sub-assemblies. Taking action a_1 costs $c_1(a_1)$. It is natural to normalize $c_1(\emptyset) = 0$ and to assume that the cost is non-decreasing in the size of the set of addressed events:

$$\text{Assumption 1:} \qquad \text{if } G_1 \subset G_2 \in \mathcal{F}, \text{ then } \; c_1(G_1) \leqslant c_1(G_2). \tag{1}$$

In period 2 (after the signal), further events can be addressed in the form of action $a_2 \in \mathcal{F}$. We assume that each event needs to be addressed only once, so we need to consider only actions a_2 "on top" of a_1: $a_1 \cap a_2 = \emptyset$. Taking action a_2 costs $c_2(a_1, a_2)$. We take again the cost of doing nothing as zero: $c_2(a_1, \emptyset) = 0$.

$$\text{Assumption 2:} \qquad \text{if } G_1 \subset G_2 \in \mathcal{F}, \text{ then } \; c_2(a_1, G_1) \leqslant c_2(a_1, G_2). \tag{2}$$

It is again natural to assume that c_2 is non-decreasing in a_2 (the more events we address in period 2, the more costly). Further, c_2 is also and non-increasing in a_1: the more events have already been addressed in period 1, the easier it is to address events in period 2. If there exists some similarity among actions, the inequality holds strictly, in absence of any similarity (3) becomes an equality:

$$\text{Assumption 3:} \qquad \text{if } G_1 \subset G_2 \in \mathcal{F}, \text{ then } \; c_2(G_1, a_2) \geq c_2(G_2, a_2). \tag{3}$$

Finally, we assume that c_2 is "bigger" than c_1, that is, later actions are more costly. This assumption avoids trivial cases, as otherwise there would be no reason at all to act in period 1 ($a_1 = \emptyset$).

$$\text{Assumption 4}: \qquad \text{for any } A, G \in \digamma,\ c_2(A, G) \geq c_1(G). \tag{4}$$

After the uncertainty has been resolved (after the event A_i containing ω has been identified), corrective action needs to be taken, or equivalently, a "shortage cost" $c_3(a_1 \cup a_2,\ A_i)$ must be paid if the actually occurred event was not addressed by either a_1 or a_2. Thus, we assume no extra "dismantling costs" at the end: $c_3(a_1 \cup a_2,\ A_i) = 0$ if $A_i \subset (a_1 \cup a_2)$. This cost structure is a generalization of the traditional cost for "wasted actions," $c_1(.)$ and $c_2(.)$ corresponding to overage costs and $c_3(.)$ to underage costs.

As before, we assume that the shortage cost is non-increasing in the set of events covered (the corrective action might benefit from having something similar to build upon) and "bigger" than the cost of addressing the event in the first place (this excludes the trivial case where it is optimal to always wait until the end):

$$\text{Assumption 5}: \qquad \text{if } G_1 \subset G_2 \in \digamma, \text{ then } c_3(G_1, A_i) \geq c_3(G_2, A_i) \quad \text{for all } A_i \in \digamma \tag{5}$$

$$\text{Assumption 6}: \qquad \text{for any } G, A_i \in \digamma,\ c_3(G, A_i) \geq c_2(G, A_i) \tag{6}$$

The decision maker's problem is to address the uncertain event with minimum expected cost. This corresponds to solving the dynamic program[1]

$$\underset{a_1 \in \digamma}{Min}\ c_1(a_1) + E_B \left[\underset{a_2 \in \digamma \setminus a_1}{Min}\ c_2(a_1, a_2) + E_{A|B}\left[c_3(a_1 \cup a_2, A)\right]\right]$$

This dynamic program builds on a long tradition of work on (discrete time) decision making under uncertainty (e.g., Denardo 1982, Ch. 6, Heyman & Sobel 1984). The second period decision $a_2(a_1, B)$ minimizes:

$$a_2(a_1, B) = \underset{G}{\operatorname{argmin}} \left\{ c_2(a_1, G) + \frac{1}{P(B)} \int_{A_i \subset B} c_3(a_1 \cup G, A_i) P(A_i) \right\}. \tag{7}$$

[1] As $\digamma$ is countable, one may equivalently write the integrals as sums.

We refer to the resulting optimal second period cost as $c_2^*(a_1, B)$. Before the signal B is observed, the period 1 decision a_1^* minimizes:

$$a_1^* = \underset{G}{\operatorname{argmin}} \left\{ c_1(G) + \int_{B \in \Psi} c_2^*(a_1, B)\, P(B) \right\}. \tag{8}$$

In this dynamic programming formulation, the preliminary character of information is captured by a less than perfect signal, i.e. Ψ being coarser than $\mathcal{F}$. However, this *preliminary nature of information has a detrimental economic impact only if costs increase over time*, i.e. if waiting for the uncertainty resolution is more expensive than acting early ($c_3(G, A) > c_2(G', A) > c_1(A)$). This is consistent with Marschak & Radner's (1972: 57) statement that evaluating an information structure is impossible without specifying the payoff structure. The fundamental trade-off posed by preliminary information is between the error from guessing and the cost of expensive delayed action.

In the presence of this trade-off, there are three basic ways in which the decision maker can react. She can either address a set of events immediately and then, in response to the signal, address a different set of events. We refer to this strategy as *adaptive*. She can also try to address as many events as possible before it is known which one will actually occur, a strategy we refer to as *hedging*. Finally, she can choose to *wait*, doing nothing until more information is available.

These three strategies can, of course, be mixed. Our main Theorem below, characterizes how the relative emphasis on these three strategies changes with the economic parameters of the problem. To state the theorem, we must first introduce two additional structural features of the costs. They are stated in the following two definitions:

Definition 1 *The cost $c'(G, A)$ is* uniformly increased *with respect to $c(G, A)$ if $c'(G, A) = \alpha\, c(G, A)$ for some $\alpha > 1$ (and uniformly decreased is defined analogously).*

Definition 2 *We say $c_2(G, A)$ has* decreasing increments *if $G_1 \subset G_2 \in \mathcal{F}$ implies for any basic event Δ: $c_2(G_1, A \cup \Delta) - c_2(G_1, A) \geq c_2(G_2, A \cup \Delta) - c_2(G_2, A)$.*

Theorem 1 *The relative importance of the three fundamental strategies hedging, adaptive, and waiting, is influenced by information and cost structure as follows:*

1. *(hedging) If $c_3(G, A)$ increases uniformly, a_1^* and $a_2^*(a_1^*, B)$ cannot decrease.*

2. *(adaptive response) If $c_2(G, A)$ decreases uniformly, and $c_2(G, A)$ and $c_3(G, A)$ have decreasing increments, a_1^* cannot increase and $a_2^*(a_1^*, B)$ cannot decrease.*

3. *(waiting) As the signal sigma-field Ψ becomes finer, a_1^* cannot increase.*

PROOF: All proofs can be found in the appendix.

The theorem illuminates general effects of information and cost structure on the fundamental strategies of hedging, adaptive response, and waiting. We see that increasing mismatch cost $c_3(G, A)$ lead to more action prior to the resolution of uncertainty (increases in a_1^* and $a_2^*(a_1^*, B)$), which corresponds to a hedging strategy. A reduction in the second period cost of action, in contrast, provides an incentive of moving actions from period 1 to period 2, leading to a more adaptive strategy. Finally, waiting becomes more beneficial with an increase in information content in the signal.

4 Specific Problem Structures

Theorem 1 provides a fundamental result showing what influences to consider in dealing with preliminary information. We now turn to four special cases of the overall optimization problem (8). They reflect more specifically the concept of preliminary information as encountered in product development and supply chain management. For each of the four special cases of the general problem (ordered newsvendor, unstructured, decomposed, and hierarchical), we derive the optimal policies and characterize the effect of preliminary information more precisely.

4.1 Ordered Sets with Distance Measures

We express the two-period newsvendor problem with forecast updates (e.g. Donohue 1996) in the framework set by Theorem 1. The set of basic events is $A_i \in \{1, 2, 3, \ldots\}$. The period 1 action is the amount of early production, $a_1 \in \{1, 2, 3, \ldots\}$, at cost c_1 per unit. Period 2 actions are $a_2 \in \{1, 2, 3, \ldots\}$, at unit cost c_2. If demand is D, the period 3 mismatch cost is $c_3 \left(D - a_1 - a_2\right)^+$ (c_3 corresponds to the classical newsvendor "underage cost"). The standard assumption (avoiding trivial cases) is $c_3 > c_2 > c_1$.

Now, the sets of possible events are *ordered* and can be described by a distribution function. Suppose that in period one, a signal y is observed about the demand x. Demand and signal have a joint density $f(x, y)$, the marginal densities are $f_1(x)$ and $f_2(y)$, and the conditional density of demand given the signal is $h\left(x|\ y\right) = f(x, y) f_2(y)$. Denote the corresponding distribution functions with the respective capital letters. Suppose that the signal y is positively correlated with demand x such that a higher signal shifts probability mass toward larger demand everywhere. This implies that $H(A|y)$ decreases in y for all A. In this special case, Theorem 1 takes the following form.

Proposition 1 (ordered): *In this newsvendor problem, it is optimal in the second period to produce* $a_2 = \left(A\left(y\right) - a_1\right)^+$, *where* $A\left(y\right)$ *is defined as*

$$H(A(y)|\ y) = \frac{c_3 - c_2}{c_3}. \tag{9}$$

Define $y^*\left(a_1\right)$ *by* $A\left(y^*\right) = a_1$. *The optimal first period production is*

$$c_1 = c_3 \Pr\left\{y \leqslant y^*\left(a_1\right);\ x > a_1\right\} + c_2\ \Pr\left\{y > y^*\left(a_1\right)\right\}. \tag{10}$$

Let us consider a special case where we can explicitly calculate the impact of signal quality and updating. As in Fisher & Raman (1996), suppose that demand and signal are characterized by a bivariate normal distribution with correlation ρ (this is only realistic if expected demand is high enough to make negative values in the model very unlikely). Suppose without loss of generality that the signal has been "translated" in such a way that $\mu_x = \mu_y = \mu$ and $\sigma_x = \sigma_y = \sigma$. Then the conditional demand distribution, given the signal y, is $\sim N(\mu + \rho[y - \mu]; \sigma\sqrt{1-\rho^2})$ (Hogg & Craig 1978, 119).

We can now transform the distribution in second period condition (9) into the standard normal:

$$\Phi\left(\frac{a_1 - \mu - \rho[y^* - \mu]}{\sigma\sqrt{1-\rho^2}}\right) = \frac{c_3 - c_2}{c_3}.$$

This implies directly that $\partial y^*/\partial a_1 = 1/\rho$. In other words, if the first period production increases by one, the signal that implies that this was the right decision goes up by $1/\rho$. A higher correlation coefficient makes the second period decision less variable (as the signal becomes more informative): $\partial A(y)/\partial y = 1/\rho$.

For $0 < \rho < 1$, we can apply the implicit function theorem to the first order condition, characterizing the change of the first period decision with the signal quality ρ. The expression of $\partial a_1/\partial \rho$ includes as a factor:

$$\frac{\partial^2 EC\left(a_1\right)}{\partial a_1 \partial \rho} = \frac{\partial y^*}{\partial \rho}\left[c_2 f_2\left(y^*\right) - c_3 \int_{x=a_1}^{\infty} f\left(x, y^*\right)\ dx\right] = 0 \tag{11}$$

(by the same argument as in the Proof of Proposition 1). Thus, we find that in this special case, the first period decision a_1 *is not influenced by the signal quality at all.* The reaction in the second period is improved, but does not shift the first period allocation. This result remains valid if demand and signal have a bivariate *log-normal* distribution with correlation ρ (avoiding the problem of possible negative demands). For the lognormal, $\partial y^*/\partial a_1 = y^*/a_1\rho$, for which (11) continues to apply.

In the extreme cases of perfect or no preliminary information, in contrast, the first period decision is affected: if the signal carries no information ($\rho = 0$), there is no second period production, and thus the first period decision must be increased to $F_2(a_1) = (c_3 - c_1)/c_3$.

If, on the other hand, the signal is perfect ($\rho = 1$), the second period reaction is capable of preventing the mismatch cost c_3 from ever being paid, and the first period decision can be reduced to $\mathcal{F}_2(a_1) = (c_2 - c_1)/c_2$.

4.2 Unstructured Problems and Hedging

The exchange of preliminary information is by no means restricted to ordered sets. Product development situations in particular are characterized by preliminary information defined on a *multi-dimensional design space*, rather than on a one-dimensional quantity space. Figure 3 provides an illustrative example, based on Terwiesch *et al.* (1999).

`Figure 3: Example of the unstructured case`

An automotive development team needs to decide which type of a heating system to include in the vehicle. There are six options available ($\mathcal{F}$ has six basic events A_i). Typically, heating system development is not on the critical path. Uncertainty about customer preferences will be resolved through market research. Marketing will release preliminary information in the form of a short list (B) by March. The final decision (A_i) is not known before October. The development group can prepare for one or more of the six outcomes by designing a system and an interface to the rest of the vehicle. If this process is started early, the only cost is internal engineering capacity ($c_1(a_1)$). If the design is delayed until March (when the short-list is available) the team reduces the risk of developing a concept that is not desired by the market. However, the delayed start requires faster execution, using external engineering capacity at higher cost ($c_2(a_2)$). If the design is delayed until October, the team will benefit from the market research and need to design only a single concept. However, this implies the risk that the heating system may delay the critical path of the overall vehicle project, causing a more than tenfold cost increase ($c_3(.,.)$).

An important element of this design problem is that the cost of covering one outcome is completely independent of covering other outcomes, i.e. covering basic event A_1 does not help in covering the occurrence of A_2. Thus, there is a specific action for each event. Then the cost functions become additive over actions:

$$c_1(a_1) = \sum_{A_i \in a_1} c_1(A_i); \quad c_2(a_1, a_2) = \sum_{A_j \in a_2} c_2(A_j); \quad c_3 = \begin{cases} 0\text{: if } A_k \in a_1 \cup a_2 \\ c_3(a_1 \cup a_2, A_k)\text{: otherwise} \end{cases}$$

The cost of covering a set of events a_2 in period 2 is independent of what events were covered in period 1. As before, costs increase over time. No cost symmetry across actions is required. In period 2, the decision maker receives a signal B, which allows updating of the outcome probabilities. The optimization problem (8) then becomes:

$$\underset{a_1 \in F}{Min}\; c_1(a_1) + E_B \left[\underset{a_2 \in F \setminus a_1}{Min}\; c_2(a_2) + E_{A|B}\left[c_3(a_1 \cup a_2, A)\right]\right]$$

Proposition 2 (unstructured): *The optimal policy for an unstructured problem is:*

$$a_1^* = \left\{ A_i \in F\colon\; c_1(A_i) \leqslant \int_B \min\left\{c_2(A_i); c_3(A_i)P(A_i|B)\right\} P(B) \right\} \tag{12}$$

$$a_2^*(a_1, B) = \left\{ A_i \in F,\; A_i \notin a_1\colon\; \frac{c_2(A_i)}{c_3(A_i)P(A_i|B)} \leqslant 1 \right\} \tag{13}$$

Consistent with Theorem 1, a higher mismatch cost increases both a_1^* and $a_2^*(a_1, B)$, and a higher second period cost shifts some actions from after the signal to prior the signal. A sharper signal reduces the cost of waiting for period 2 until taking action.

For the heating system development example, this result has several important consequences. First, the higher the time pressure on the overall project (the cost of delaying the critical path), the more actions one should move forward in time, even if they are potentially wasted. We refer to this approach as *hedging*, as it protects the decision maker from incurring the mismatch cost. Second, an increase in $c_2(A_k)$ leads to more action prior to the signal. Finally, the fewer alternatives are included in the "short-list", the less action should be carried out in period 1. A highly informative signal allows reacting more effectively in period 2, reducing the risk of falling short.

4.3 Decomposable Problem and Pipelining

Complex decisions, especially in the field of product development, can frequently be broken up into two (or more) smaller problems that are tackled in parallel or in series. To capture such decomposable problems, consider two separate domains, $-_{\alpha}$ and $-_{\beta}$, each having their respective sigma fields of distinguishable events. Events in $-$ are now a combination $(A_i^{\alpha}, A_j^{\beta}) \in \mathcal{F}_{\alpha} \times \mathcal{F}_{\beta} = \mathcal{F}_{\alpha\times\beta}$. As in the general case, the decision maker can choose actions $a_t^{\alpha} \in \mathcal{F}_{\alpha}$ and $a_t^{\beta} \in \mathcal{F}_{\beta}$ $(t = 1, 2)$, based either on her prior $(t = 1)$ or on preliminary information $(t = 2)$.

Figure 4: Example of the decomposable case

Figure 4 illustrates the concept of decomposability. Consider a development team that evaluates six suppliers for two components, a control unit with three potential suppliers (defining $\mathcal{F}_{\alpha}$) and a ventilation unit with three different suppliers (defining $\mathcal{F}_{\beta}$). Suppose a supplier evaluation team with scarce capacity performs the selection sequentially, evaluating the supplier for control unit first. The control unit supplier selection can be released early (in form of B_{α}) but carries no information about the best supplier choice for the ventilation unit.

The chosen supplier installs a first testing tool-set in the plant to fine-tune its performance. As tool installation is on the critical path of the overall development project for both components, the team considers tool installation in parallel to supplier selection. An obvious improvement is to start with the control unit tool as soon as its supplier is selected (rather than wait for the selection of the ventilation unit supplier). This method is referred to as "task-pipelining" (Ulrich & Eppinger 1999, Krishnan *et al.* 1997[2]). This rule of thumb assumes that the α-uncertainty is *completely* resolved in period 2, which is often

[2]The concept of pipelining can also be found as a command in the operating system UNIX, where a "|" between two commands allows an early start of the second command.

not the case.

In contrast to a hierarchical problem structure (considered in the next section), the decisions in the two domains are *independent*: a decision in the $\alpha-$domain has no impact on the $\beta-$domain and vice-versa. This allows the decision maker to choose actions from $\mathcal{F}_\alpha$ and $\mathcal{F}_\beta$ rather than from $\mathcal{F}_{\alpha\times\beta}$[3].

The uncertainty of one domain, say, domain α, is reduced first (otherwise, we are back in Section 4.2). Preliminary information arises in form of a signal $B_\alpha \in \Psi_\alpha \subset \mathcal{F}_\alpha$. Define:

$$\begin{aligned} c_t^\alpha(a_t^\alpha) &= \sum_{A^\alpha \in a_t^\alpha} c_t^\alpha(A^\alpha); \qquad c_t^\beta(a_t^\beta) = \sum_{A^\beta \in a_t^\beta} c_t^\beta(A^\beta), \qquad t = 1, 2 \\ c_t(a_t) &= c_t^\alpha(a_t^\alpha) + c_t^\beta(a_t^\beta) \\ c_3(a_1 \cup a_2, A) &= c_3^\alpha(a_1^\alpha \cup a_2^\alpha, A^\alpha) + c_3^\beta(a_1^\beta \cup a_2^\beta, A^\beta) \end{aligned}$$

Costs are again additive over events and actions, as well as over the two design domains, $\mathcal{F}_\alpha$ and $\mathcal{F}_\beta$, due to their independence. This creates the overall problem:

$$\underset{a_1^x \in \mathcal{F}_x}{Min} \sum_{x=\alpha,\beta} c_1^x(a_1^x) + E_{B_\alpha}\left[\underset{a_2^x \in \mathcal{F}_x}{Min} \sum_{x=\alpha,\beta} c_2^x(a_2^x) + E_{x|\,B_\alpha} \sum_{x=\alpha,\beta} c_3^x(a_1^x \cup a_2^x, A^x)\right] \tag{14}$$

In addition to a separability of the cost functions, decomposability also leads to an independence of the probability structure: the signal B_α carries no information about the β-event A^β. Thus, $E_{\beta|\,B_\alpha}[c_3^\beta(a_1^\beta \cup a_2^\beta, A^\beta)] = \sum_{A^\beta} c_3^\beta(a_1^\beta \cup a_2^\beta, A^\beta)\ P(A^\beta)$.

Proposition 3 (decomposable): *The optimal policy for a decomposable problem is:*

$$\begin{aligned} a_1^{\alpha *} &= \left\{A^\alpha \in \mathcal{F}_\alpha\colon\ c_1^\alpha(A^\alpha) \leqslant \int_{B_\alpha} \min\left\{c_2^\alpha(A^\alpha); c_3^\alpha(A^\alpha) P(A^\alpha|B_\alpha)\right\} P(B_\alpha)\right\}; \\ a_2^{\alpha *}(a_1^\alpha, B_\alpha) &= \left\{A^\alpha \in \mathcal{F}_\alpha \setminus a_1^\alpha\colon\ \frac{c_2^\alpha(A^\alpha)}{c_3^\alpha(A^\alpha) P(A^\alpha|B_\alpha)} \leqslant 1\right\}. \\ a_1^{\beta *} &= \left\{A^\beta \in \mathcal{F}_\beta :\ \frac{c_1^\beta(A^\beta)}{c_3^\beta(A^\beta) P(A^\beta)} \leqslant 1\right\}; \qquad a_2^{\beta *}(a_1^\beta, B_\alpha) = \emptyset. \end{aligned}$$

[3] This simplifies the problem substantially. Consider the case in which $\mathcal{F}_\alpha$ has 5 and $\mathcal{F}_\beta$ 7 elements. Choosing from $\mathcal{F}_{\alpha\times\beta}$ would require a consideration of 5*7=35 elements. Choosing from $\mathcal{F}_\alpha$ and $\mathcal{F}_\beta$ directly, though, only considers 5+7=12 elements (see Simon 1969, 206).

Proposition 3 follows directly from treating both domains as separate unstructured problems. An $\alpha-$event A^α may best be addressed before or after the signal or after complete uncertainty resolution, as in Proposition 2. The $\beta-$action, in contrast, has no signal (B_α carries no information about A^β), which makes any period 2 action senseless. Thus, $\beta-$action should be carried out either immediately or after final uncertainty resolution.

Proposition 3 generalizes the idea of pipelining from a deterministic to a stochastic setting. "Stochastic pipelining" has three important implications. First, the information concerning the $\alpha-$domain need not be 100% reliable: depending on the cost structure, even a weaker signal may allow capturing some pipelining benefits. Second, stochastic pipelining can include an element of hedging (similar to the unstructured case): it is not the signal alone that determines when it is optimal to execute an action, but a combination of the signal and the cost structure (as we discuss following Proposition 2). Third, hedging in the $\beta-$domain may be optimal even without a signal, if early action is sufficiently cheap and prior knowledge sufficiently sharp.

4.4 Hierarchical Problem Structures

A hierarchical problem involves again two types of actions and information, represented by basic events in an $\alpha-$domain and in a $\beta-$domain. However, the two domains are now *dependent*. Let the $\beta-$decision be *sequentially dependent* on the $\alpha-$decision. This has an important consequence: while in the independent case, actions were chosen from either $\mathcal{F}_\alpha$ or $\mathcal{F}_\beta$, the choice A^β now "assumes" an element in the $\alpha-$domain.

`Figure 5: Example of the hierarchical case`

This situation is illustrated by the example in Figure 5. A manufacturer of semiconductor production equipment is expecting a customer order. The company produces two different platforms, A and B, which are configured and assembled to detailed customer specifications. At the current time, it is uncertain which of the two platforms the customer will pick. No firm purchase order has yet been signed, and thus the customer may not order at all. The purchase order includes a specification of the platform as well as a delivery date (the α-domain includes 'no order', 'platform 1', and 'platform 2').

Each of the two platforms can be configured in three different ways, numbered 1-3 for A and 4-6 for B. Thus, $A^\beta \in \{1, \ldots, 6\}$. No components are shared across platforms, i.e., building any given component "assumes" its platform (for example, building component 5 assumes that platform 2 has been ordered).

Despite significant lead-times for placing orders to their suppliers, the manufacturer works without any sub-assemblies in inventory (as a result of high holding cost, due to expensive components and fast obsolescence). Once the platform is known, the platform specific components can be ordered (a_t^β is placed). In the spirit of postponement (see e.g. Lee & Tang 1997), it is possible to delay orders for the specific configuration until the full specifications are available.

All orders placed in the first period (prior to receiving the purchase order for the platform) can follow the standard procurement process (leading to c_1). If an order is placed later, it has to be expedited, leading to increased shipping cost and a premium charge from the supplier (c_2). If the components are not procured on time, the promised shipment date is at risk, leading to severe penalty charges (c_3).

Sequential dependence allows us to define a set function $f : \digamma_\alpha \rightarrow \digamma_\beta$, with $f(A^\alpha)$ defining the set of $\beta-$decisions compatible with a given A^α. In the example, $f(A) = \{1, 2, 3\}$ and $f(B) = \{4, 5, 6\}$. Conversely, call $f^{-1}\left(A^\beta\right) = \left\{A^\alpha\text{: } A^\beta \in f(A^\alpha)\right\}$ the set of α-decisions implying A^β. If there is no component commonality each $f^{-1}\left(A^\beta\right)$ has only one element: in the example, $f^{-1}(5) = B$.

The joint prior probability measure has the structure $P\left(A^\alpha, A^\beta\right) = 0$ whenever $A^\beta \notin f(A^\alpha)$. The marginal probabilities become $P(A^\alpha) = \sum_{A^\beta \in f(A^\alpha)} P\left(A^\alpha, A^\beta\right)$, and $P(A^\beta) = \sum_{A^\alpha \in f^{-1}(A^\beta)} P\left(A^\alpha, A^\beta\right)$. As in the decomposable case, uncertainty related to α is reduced by a signal $B \in \Psi_\alpha$. For example, the customer commits to an order, with penalties in case of a cancellation, but does not yet indicate the platform. The signal $B \in \Psi_\alpha$ carries less information for the β-choice: while $P(A^\alpha |\ B) = P\left(A^\alpha\right)/P(B)$ if $A^\alpha \subset B$ and zero otherwise, we must write $P(A^\beta |\ B) = \sum_{A^\alpha \subset B \cap f^{-1}(A^\beta)} P\left(A^\alpha, A^\beta\right)/P(B)$.

The decision maker may postpone the $\beta-$decision, a_1^β, until more information is available. After the signal B, one needs to consider only $\beta-$choices in $\cup_{A^\alpha \subset B} f(A^\alpha)$, or if one even waits until period 3, both A^α and $A^\beta \in f(A^\alpha)$ are known. Alternatively, the decision maker may also choose a_1^β before a signal is received. This might be beneficial if one has sufficient "hope" that the outcome will confirm the chosen a_1^β.

Suppose for now that costs $c_t^\alpha(a_t^\alpha)$ and $c_t^\beta(a_t^\beta)$ are additive over the two domains. That is, making an upstream choice a_1^α does not reduce the component costs c_t^β. This is realistic in a configuration (as opposed to a new product development) problem, where the components are known, and the question is picking the right combination of them. The resulting optimization problem becomes:

$$\underset{x \in \{\alpha,\beta\}; a_1^x \in F^x}{Min} \sum_{x=\alpha,\beta} \left\{ c_1^x(a_1^x) + E_B \left[\underset{a_2^x \in F^x \setminus a_1^x}{Min} \; c_2^x(a_2^x) + E_{x|B}\left[c_3^x(a_1^x \cup a_2^x, A^x)\right] \right] \right\}$$

This problem can be decomposed into α- and β parts as long as the costs are additive across domains. Each corresponds to the unstructured problem analyzed earlier.

In general, however, we have to allow a certain downstream $\beta-$decision to be common *across* A^α's $(f(A_1^\alpha) \cap f(A_2^\alpha) \neq 0)$, thereby leading to a sub-additive cost structure. There are also important cases with synergies among multiple $\beta-$decisions for a *given* upstream event α. To understand how sub-additive cost structures influence the management of preliminary information, we define the following.

Definition: *A problem exhibits* downstream commonality *if downstream actions are common across upstream events: across a set of upstream events* $\{A_1^\alpha, \ldots, A_n^\alpha\}$ *,there are downstream events* $A_i^\beta \in f(A_i^\alpha)$ *for all* i*, such that* $c_t^\beta \left(\bigcup_i (A_i^\beta)\right) < \sum_i c_t^\beta (A_i^\beta)$, $t = 1, 2$ *(see assumption 3). If there is an* $\bar{A}^\beta \in f(A^\alpha)$ *for all* A^α *(it needs to be chosen only once for all possible upstream events) we refer to this as a universal downstream action.*

Definition: *A problem exhibits* upstream synergy *if downstream actions associated with an upstream event have decreasing increments: there is an* A^α *with a* β*-set* $G(A^\alpha) \subset f(A^\alpha)$ *such that* $c_t^\beta \left(\bigcup_{A^\beta \in G(A^\alpha)} A^\beta\right) < \sum_{A^\beta \in G(A^\alpha)} c_t^\beta (A^\beta)$, $t = 1, 2$*. If costs are incurred only for the first downstream action* $\bar{A}^\beta$*, we refer to this as perfect synergy.*

In both cases, the subadditivity applies only to periods 1 and 2 – the period 3 cost is a "market penalty" defined for one event that does occur without having been covered.

`Insert Figure 6: Component commonality and platform synergy`

Downstream commonality may make it attractive to take a downstream action prior to the signal, despite the sequential dependence. Consider the case in which a new machine needs to be installed in the factory (fab), including configuration and shipment of the machine (upstream) and connecting the machine to various support systems, e.g. computer networks, chemical supply pipes, and transportation devices (downstream). Downstream commonality in this case implies that some of the activities related to the support systems (e.g. installing the supply pipes) are independent of the configuration. One way to cut the overall installation time is to build a *template* of the actual machine, an empty box with the same external geometry and interfaces, and to connect this template to the rest of the fab (thereby addressing a downstream event $\bar{A}^\beta$ in period 1). Once the actual machine arrives, it simply replaces the template, which is discarded or used for other installations.

Upstream synergy exists, if the sub-additive costs result within one upstream solution. For example, Lee and Tang (1997) describe the case of a printer with a common platform for various configurations. This may make it optimal to *postpone* certain downstream activities until the detailed configuration is known (address all downstream events corresponding to upstream event $\bar{A}^{\alpha}$ after $\bar{A}^{\alpha}$ is confirmed). Templating can be thought of as being the "opposite" of postponement. Postponement starts upstream and delays the downstream task, templating starts downstream before upstream.

Proposition 4 (hierarchical): *(a) If the configuration space exhibits downstream commonality, then a_2^{β}, the set of events A^{β} covered in the second period, is as large or larger compared to the situation without downstream commonality.*

(b) If one downstream event $\bar{A}^{\beta}$ is universal, a sufficient condition making templating of $\bar{A}^{\beta}$ optimal is: $c_1\left(\bar{A}^{\beta}\right) \leqslant c_3\left(\bar{A}^{\beta}\right) P\left(\bar{A}^{\beta}\right)$.

(c) If an upstream event $\bar{A}^{\alpha}$ exhibits upstream synergy, then a_2^{β}, the set of events $A^{\beta} \in f\left(\bar{A}^{\alpha}\right)$ covered in the second period, is as large or larger compared to the situation without upstream synergy.

(d) If an upstream event $\bar{A}^{\alpha}$ exhibits perfect synergy, a sufficient condition for optimal postponement is: $c_2^{\beta}\left(A^{\beta}\right) > \int_{A^{\beta} \in f\left(\bar{A}^{\alpha}\right)} c_3^{\beta}\left(A^{\beta}\right) P\left(\bar{A}^{\alpha} \mid B\right)$ and $c_1^{\beta}\left(A^{\beta}\right) > c_2^{\beta}\left(A^{\beta}\right) P\left(\bar{A}^{\alpha}\right)$.

5 Conclusion

Decision makers today rarely have the luxury of delaying their decisions until all of the required information input has become available. Rather than waiting, it is often preferable to make early decisions based on preliminary information and then adjust these decisions as more information arrives. This trend towards using preliminary information has been accelerated by recent advances in communication technologies. With the marginal cost of information exchange approaching zero, it becomes economically beneficial to exchange and update information earlier than before.

This paper proposes a generic model of the choices under preliminary information: early decisions at low cost but at a high risk of doing the "wrong thing," versus better informed decisions at higher cost. We describe preliminary information in the form of a sigma field that is refined over time. We are able to show several properties of optimally using preliminary information under very general assumptions. Increasing mismatch costs lead to more action prior to the resolution of uncertainty, a strategy we refer to as hedging. A reduction in the second period cost of action, in contrast, leads to a more adaptive strategy. An increase in information content in the signal, leads to a delay in action. This fundamental result (Theorem 1) provides the answer to the question in our title ("Rush and Be Wrong or Wait and Be Late?").

Preliminary information is common across applications in product development and supply chain management. We derive optimal policies for four special instances of the general problem, ordered (newsvendor), unstructured, decomposed, and hierarchical. These optimal policies correspond to specific strategies that have been discussed in literature and applications. *Hedging* describes a strategy that starts several actions simultaneously, knowing that only one of them will ultimately be used. Stochastic *pipelining* can be used if the incoming information can be decomposed, allowing for an early start of the first set of actions. *Postponement* delays downstream actions related to information that is anticipated to come in late, while starting upstream actions early that are common across scenarios. Finally, *templating* allows scenario-common downstream actions to start early (at low cost), using a mock-up presentation of some upstream action. Our results generalize these strategies to situations where they apply only "imperfectly", or stochastically.

Our model opens a new way of analyzing operations management decisions in response to preliminary information. Several previous studies have recognized the importance of information sharing in product development, but none has represented preliminary information except via a one-dimensional parameter with a distribution function. In the field of supply chain management, previous studies have focused primarily on newsvendor-type settings, avoiding more complex decision space topologies.

To conclude, we propose three promising directions of future research that build directly on the findings presented here. First, the signal providing preliminary information is often not exogenous, but rather the result of a costly action. For example, Fisher & Raman's (1996) apparel company has to produce early in the season to get the benefit of the more accurate early seasons sales forecast. Both our general model and the specific problem structures can be extended to account for such dependence.

Second, our model assumes that each possible outcome state defines exactly one action, over the set of which the decision maker chooses. However, there may be situations with "flexible" actions. Uncertainty about the components required, the company needs to decide what orders to place. If there are two possible outcomes for a specific component (required / not required), the *ex-post* optimal decisions are to either order or not order. However, there might be a third possible action, to purchase an option for the component (e.g., pay a fee to reserve a capacity slot). While this decision will never be optimal *ex-post*, it may well be beneficial *ex-ante* (see, e.g., van Mieghem 1998). Such flexible actions are not included in our current model formulation.

Finally, the wasted cost of addressing an event that ultimately does not occur can be larger that just the cost of the action. For example, if a maker of stamping dies cuts off parts of the metal as a response to preliminary information, this does not only cost time and effort, but might require a scrapping of the whole die if the information turns out to be wrong. Including a cost of "undoing" an action may change the results of our model.

Both the product development literature and the supply chain literature list the management and exchange of information as a research priority for the coming years. Yet, there exists almost no overlap in citations or methodology across the two fields. We hope that the model presented here can help to bring together these streams of research. Recognizing structural similarities of the respective decision problems will allow a more general analysis of the management of preliminary information in the future.

References

Cachon, G., "Competitive Supply Chain Inventory Management", in Tayur, S., R. Ganeshan, M. Magazine, *Quantitative Models for Supply Chain Management*, Kluwer 1999, 111 - 146.

Clark, K. B., and T. Fujimoto. "Lead Time in Automobile Development: Explaining the Japanese Advantage", *Journal of Technology and Engineering Management* 6, 1989, 25-58.

Cohen, M., T. Ho, J. Ren, C. Terwiesch. "Cycle Time Performance in the Semiconductor Equipment Industry", presented at INTEL University Day, Santa Clara, 2000.

Denardo, E. V., *Dynamic Programming: Theory and Applications*, Englewood Cliffs: Prentice Hall 1982.

Donohue, K. L., "Supply Contracts for Fashion Goods: Optimizing Channel Profits", Working Paper, The Wharton School, 1996.

Eppen, G. D. and A. V. Iyer, "Improved Fashion Buying Using Bayesian Updates", *Operations Research* 45 (6), 1997, 805 - 819.

Fisher, M. and A. Raman, "Reducing the cost of demand uncertainty through accurate response to early sales", *Operations Research* 44 (1), 1996, 87 - 99.

Hammond, J., "Quick Response in the Apparel Industry", HBS Note N9-690-038, Harvard Business School Publishing, Boston, 1990.

Heyman, D. P. and M. J. Sobel, *Stochastic Models in OR* Vol II, New York: McGrawHill 1984.

Hogg, R. V. and A. T. Craig. *Introduction to Mathematical Statistics*. 4th ed., New York: McMillan 1978.

Krishnan, V., S. D. Eppinger, and D. E. Whitney, "A Model-Based Framework to Overlap Product Development Activities", *Management Science* 43, No. 4, 1997, 437-451.

Lee, H. L. and C. S. Tang, "Modeling the Costs and Benefits of Delayed Product Differentiation", *Management Science* 43 (1), 1997, 40 - 53.

Loch, C. H. and C. Terwiesch, "Communication and Uncertainty in Concurrent Engineering", *Management Science* 44 (8), 1998, 1032 - 1048.

Marschak, J. and R. Radner. *Economic Theory of Teams*, New Haven: Yale University Press 1972.

McCardle, K. F. "Information Acquisition and the Adoption of New Technology," *Management Science* 31, 1985, 1372 - 1389.

McCardle, K. F. and S. A. Lippman "Does Cheaper, Faster or Better Imply Sooner in the Timing of Innovation Decisions?" *Management Science* 33, 1987, 1058 - 1064.

Nahmias, S. *Production and Operations Analysis*, 2nd ed., Homewood IL: Irwin, 1993.

Simon, H. A., *The Science of the Artificial*, Cambridge, MA: MIT Press 1969.

Terwiesch, C., C. H. Loch and A. De Meyer. "Exchanging Preliminary Information in Concurrent Development Processes", Working Paper Wharton and INSEAD, 2000

Ulrich, K. and S. Eppinger. *Product Design and Development*, New York: McGraw-Hill, 1994.

Van Mieghem, J. A. "Investment Strategies for Flexible Resources," *Management Science* 44 (8), 1998, 1071 - 1078.

6 Appendix

Proof of Theorem 1. To prove claim (1), we rewrite (8) more explicitly as: a_1^* minimizes over possible x:

$$c_1(x)+\int_{B\in\Psi}\left[c_2\left(x,a_2\left(x,B\right)\right)+\frac{1}{P\left(B\right)}\int_{A_i\subset B}\alpha\ c_3(x\cup a_2\left(x,B\right),A_i)P\left(A_i\right)\right]P\left(B\right).$$

Because of the multiplication of c_3 by α, the value of adding some A_j to $(x \cup a_2)$ increases. Therefore, the optimal x will not decrease and, as this (weakly) decreases the second period costs and its increments (assumption 3 and decreasing increments), $a_2\left(a_1^*, B\right)$ does not decrease (or possibly increases). This proves claim (1).

To prove claim (2), we rewrite (8) again, this time with the new period 2 costs. a_1^* minimizes:

$$c_1(x)+\int_{B\in\Psi}\left[\alpha\ c_2\left(x, a_2\left(x, B\right)\right)+\frac{1}{P\left(B\right)}\int_{A_i\subset B} c_3(x\cup a_2\left(x, B\right), A_i)P\left(A_i\right)\right]P\left(B\right).$$

The multiplication of c_2 by α increases the benefit both of reducing a_2 or increasing x. If a_2 is reduced, $c_3(x\cup a_2, A_i)$ grows, and thus by decreasing increments, the value of increasing x increases. Thus, the optimal x does not shrink. If, in turn, x is increased, the benefit of reducing a_2 grows because of decreasing increments (it reduces c_2 more and increases c_3 less). Thus, the optimal a_2 does not grow. This proves claim (2).

Turning to claim (3), observe first that $c_2^*\left(a_1, B\right)$, the second period cost that is minimized by $a_2(a_1, B)$, is non-increasing in a_1 because both cost arguments weakly decrease in a_1. Observe also the following lemmas:

Lemma 1. If $\Psi_1\subset\Psi_2$, $\int_{B\in\Psi_1} c_2^*\left(a_1, B\right)P\left(B\right)\geq\int_{G\in\Psi_2} c_2^*\left(a_1, G\right)P\left(G\right).$

Proof of Lemma 1. $\Psi_1\subset\Psi_2$ means that any $B\in\Psi_1$ fulfills $B=G_1\cup G_2$, both in Ψ_2 (one of the two may be the null set). Suppose WLOG that all Ψ_1 sets and Ψ_2 sets are the same, $B^{'}=G^{'}$, except for *one* $B=G_1\cup G_2$. The for all other sets $B^{'}$, $c_2^*\left(a_1, B^{'}\right)=c_2^*\left(a_1, G^{'}\right)$. But on set B, by the fact that one can optimize on G_1 and G_2 separately:

$$\begin{aligned}
c_2^*\left(a_1, B\right) &= \min_x\left\{c_2(a_1, x)+\frac{1}{P\left(B\right)}\int_{A_i\subset B} c_3(a_1\cup x, A_i)P\left(A_i\right)\right\}\\
&\geq \sum_{j=1,2}\frac{P(G_j)}{P\left(B\right)}\min_{y_j}\left\{c_2(a_1, y_j)+\frac{1}{P\left(G_j\right)}\int_{A_i\subset G_j} c_3(a_1\cup y_j, A_i)P\left(A_i\right)\right\}\\
&= \int_{G=G_1,G_2} c_2^*\left(a_1, G\right)P(G). \qquad \text{This proves Lemma 1.}\square
\end{aligned}$$

Lemma 2. An increase in a_1 has a lower benefit if the signal comes from a finer sigma field: If $\Psi_1 \subset \Psi_2$, then for all $\Delta \in F$:

$\int_{B\in\Psi_1}[c_2^*(a_1,B) - c_2^*(a_1\cup\Delta,B)]P(B) \geq \int_{G\in\Psi_2}[c_2^*(a_1,G) - c_2^*(a_1\cup\Delta,G)]P(G)$.

Proof of Lemma 2. We can rearrange the terms as $\int_{B\in\Psi_1} c_2^*(a_1,B)\,P(B) - \int_{G\in\Psi_2} c_2^*(a_1,G)\,P(G) \geq$

$\int_{B\in\Psi_1} c_2^*(a_1\cup\Delta,B)\,P(B) - \int_{G\in\Psi_2} c_2^*(a_1\cup\Delta,G)\,P(G)$, and we know both sides of this expression are positive by Lemma 1. The fact that the right hand side of this inequality is smaller follows from the assumption of decreasing increments of $c_2(\cdot)$ (the algebraic details are omitted). This proves Lemma 2.□

The Theorem follows directly. From (8) and Lemma 2, the benefit of increasing a_1 is smaller in the finer sigma field Ψ_2, and therefore a_1^* can only decrease and not increase.□

Proof of Proposition 1. The expected cost period 2 cost, given a_1 and y, is

$$EC(a_2) = c_2a_2 + c_3\int_{z=a_1+a_2}^{\infty}(z-a_1-a_2)\;h(z|\;y)\,f(z)\,dz,$$

which can be differentiated with respect to a_2, yielding (9) as a first order condition. It can easily be checked that $EC(a_2)$ is convex. The solution $A(y)$ is unique (the numerator grows in $A(y)$ while the right hand side is constant and < 1). Furthermore, the first order condition (9) together with the assumption that $H(A|y)$ decreases in y implies that $A(y)$ increases in y. This, in turn, implies that $y^*(a_1)$ increases in a_1. The first period cost resulting from choosing a_1 is:

$$\begin{aligned} EC(a_1) &= c_1a_1 + c_3\int_{y=0}^{\infty}\int_{x=Max\{a_1,A(y)\}}^{\infty}(x - Max\{a_1,A(y)\})\;h(y|\;x)\;f_2(y)\,dx\;dy \\ &\quad + c_2\int_{y=y^*(a_1)}^{\infty}[A(y) - a_1]\;f_2(y)\;dy. \end{aligned}$$

The first derivative of $EC(a_1)$ is $c_1 - c_3\int_{y=0}^{y^*(a_1)}\int_{x=a_1}^{\infty} f(x,y)\;dx\;dy - c_2\int_{y=y^*(a_1)}^{\infty} f_2(y)\;dy$, which can be written as (10) when set to zero. The second derivative shows that the first period cost is convex:

$$
\begin{aligned}
\frac{\partial^2 EC\left(a_1\right)}{\partial a_1^2} &= c_3 \int_{y=0}^{y^*} f\left(a_1, y\right)\ dy + \frac{\partial y^*}{\partial a_1}\left[c_2 f_2\left(y^*\right) - c_3 \int_{x=a_1}^{\infty} f\left(x, y^*\right)\ dx\right] \\
&= c_3 \int_{y=0}^{y^*} f\left(a_1, y\right)\ dy + c_3 \frac{\partial y^*}{\partial a_1}\left[\frac{c_2 - c_3}{c_3} f_2\left(y^*\right) + \int_{x=0}^{a_1} f\left(x, y^*\right)\ dx\right] \\
\text{(by 9)} \quad &= c_3 \int_{y=0}^{y^*} f\left(a_1, y\right)\ dy + c_3 \frac{\partial y^*}{\partial a_1}\left[-\int_{x=0}^{a_1} h\left(x \middle|\ y^*\right) f_2\left(y^*\right)\ dx + \int_{x=0}^{a_1} f\left(x, y^*\right)\ dx\right] \\
&= c_3 \int_{y=0}^{y^*} f\left(a_1, y\right)\ dy > 0. \ \square
\end{aligned}
$$

Proof of Proposition 2. To find the optimal policy for this decision problem, we first consider the situation after having received the preliminary information. The expected cost for choosing a_2 given a_1 and B can be written as:

$$EC(a_2, B) = \sum_{A_i \in a_2} c_2(A_i) + \frac{1}{P(B)} \int_{A_j \in B \cap a_2^c} c_3(a_1 \cup a_2, A_j) P(B)$$

including the cost of action and the updated expectation of the action falling short. For any event $A_k \notin a_1$, adding A_k to a_2 will change the expected cost $EC(a_2)$:

$$\Delta EC : c_2(A_k) - \frac{1}{P(B)} c_3(A_k) P(A_k) \overset{(!)}{\lessgtr} 0 \Longleftrightarrow \frac{c_2(A_k)}{c_3(A_k) P(A_k|B)} \overset{(!)}{\lessgtr} 1$$

which defines the optimal second period policy $a_2^*(a_1, B)$ as stated above.

The first period cost include three elements, the cost of action in the first period, the cost action in the second period that will be invested for certain signals B and the expected cost of mismatch, given $a_2^*(a_1, B)$:

$$EC(a_1) = \sum_{A_i \in a_1} c_1(A_i) + \int_B \left[\sum_{A_j \in a_2^*(a_1, B)} c_2(A_j) + \frac{1}{P(B)} \sum_{A_j \in B, A_j \notin a_2^*(a_1, B)} c_3(A_j) P(A_j)\right] P(B)$$

Again, we look at the cost of adding A_k, this time to a_1:

$$\Delta EC : c_1(A_k) - \int_B \left[I_{\{A_k \text{ chosen in period } 2\}} c_2(A_k) + (1 - I_{\{A_k \text{ chosen in period } 2\}}) \frac{c_3(A_k) P(A_k)}{P(B)}\right] P(B)$$

The first term in the integral captures the expected cost savings of moving A_k from period 1 to period 2 and the second term captures the savings in expected mismatch cost. $\square$

Proof of Proposition 3. For each domain separately, the problem can be treated as a special case of Proposition 2. For the $\beta-$domain, the signal contains no information because the domains are independent. $\square$

Proof of Proposition 4. Consider downstream commonality first. The condition for choosing a downstream event in the second period is given by (13). Subadditivity of the cost c_2 implies that the numerator decreases, rendering the condition less stringent and possibly leading to an increase of a_2. We note here that the set of β-events addressed in period 1 may increase or decrease, depending on whether c_1 or c_2 exhibit stronger subadditivity (see Equation 12).

If a downstream event $\bar{A}^\beta$ exhibits perfect commonality, its cost is zero in period 2 if it has been included in period 1. Thus, in period 1, condition (12) becomes $c_1^\beta(\bar{A}^\beta) \leqslant \int_B \min\left\{c_2^\beta(\bar{A}^\beta); c_3^\beta(\bar{A}^\beta)P(\bar{A}^\beta \mid B)\right\} P(B)$. To consider a sufficient condition, suppose that c_3 is always the minimum in the integral (c_2^β is by Assumption 4 larger than c_1^β; taking $c_3^\beta \; P(\bar{A}^\beta \mid B)$ as the minimum is, therefore, stricter). As $f^{-1}\left(\bar{A}^\beta\right) = \text{-}\ {}_\alpha$, this becomes $c_1^\beta(\bar{A}^\beta) \leqslant c_3^\beta(\bar{A}^\beta)P(\bar{A}^\beta)$.

Next, consider upstream synergy. Again, the condition for a_2^β is given by (13). Subadditivity of the cost c_2 decreases the numerator, rendering the condition less stringent.

In the special case where the synergy is perfect on upstream event $\bar{A}^\alpha$, the optimal policy is as follows. If this first component has been produced in period 1, the optimal policy in period 2 is to produce all others. If it has not been produced in period 1, production in period 2 is optimal if its cost is lower than the expected mismatch cost, according to (13). Thus, it is optimal to wait if the first condition in the proposition holds. Immediate production of the first component, thus, avoids all period 2 costs and is worthwhile if it is cheaper than the expected mismatch cost. This corresponds to the second condition. $\square$

	Prior Information	*Action 1*	*Signal (preliminary Information)*	*Action 2*	*Resolution*	*Mismatch: Underage, Overage*	Literature
Classical News-vendor	Past demand distributions	Production	-	-	Real demand observed	Lost profit, excess inventory	Nahmias 1993
News-vendor with updating, QR	Past demand distributions	Low cost first period production	Market research, early sales	Higher cost second period production	Real demand observed	Lost profit / excess inventory	Eppen & Iyer 1997; Donohue 1997; Fisher & Anand 1996
Custom config-uration / build-to-specs	Typical customer Specification (experience)	Component procurement & capacity slotting	Customer purchase intent	Additional procure-ment, rush orders	Purchase order with final Specific-ation	Delivery delay, excess procurement	Cohen et al. 2000
Parallel activities/ Concurrent Engineering	Old designs of previous products	Early / parallel start of down-stream task	Preliminary drawing release, shared CAx files	Parallel develop-ment	Final design release	Project delay, rework and wasted effort	Clark & Fujimoto 1991; Krishnan et al. 1997; Loch & Terwiesch 1998

Figure 1: Making decisions under preliminary information

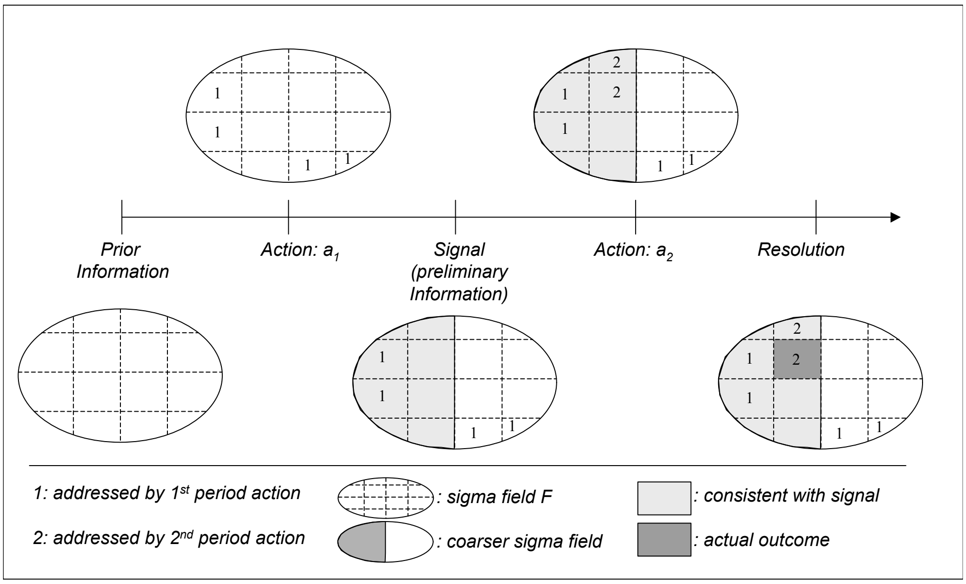

Figure 2: The general model

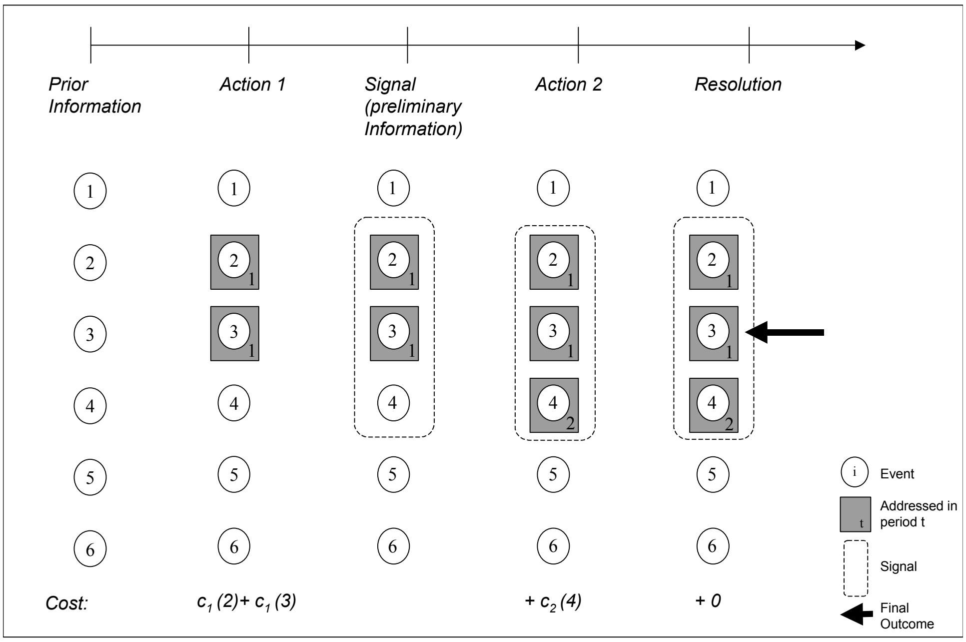

Figure 3: Example of the unstructured case

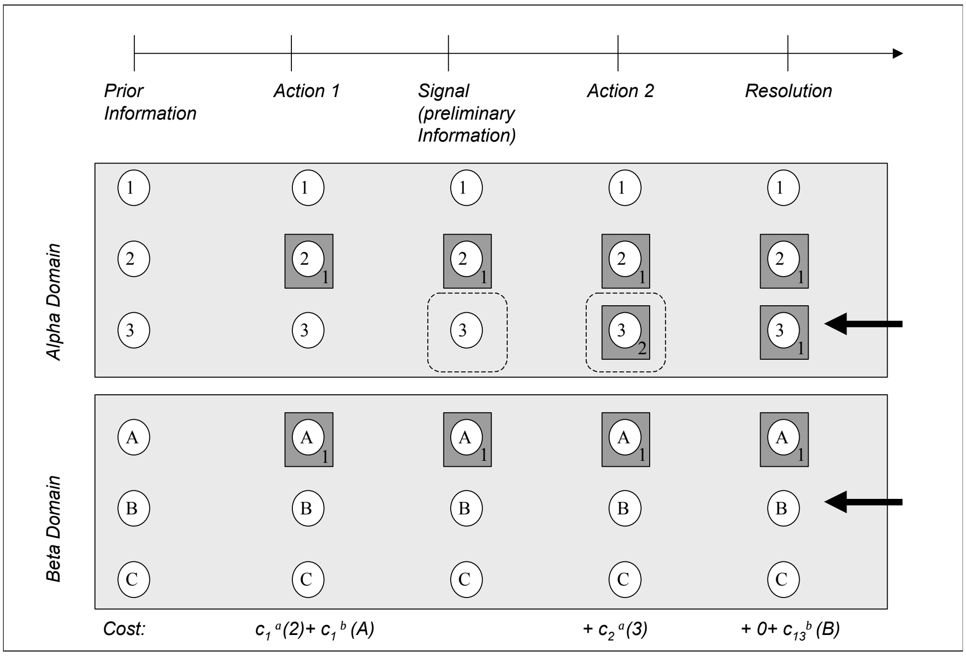

Figure 4: Example of the decomposable case (Legend see Figure 3)

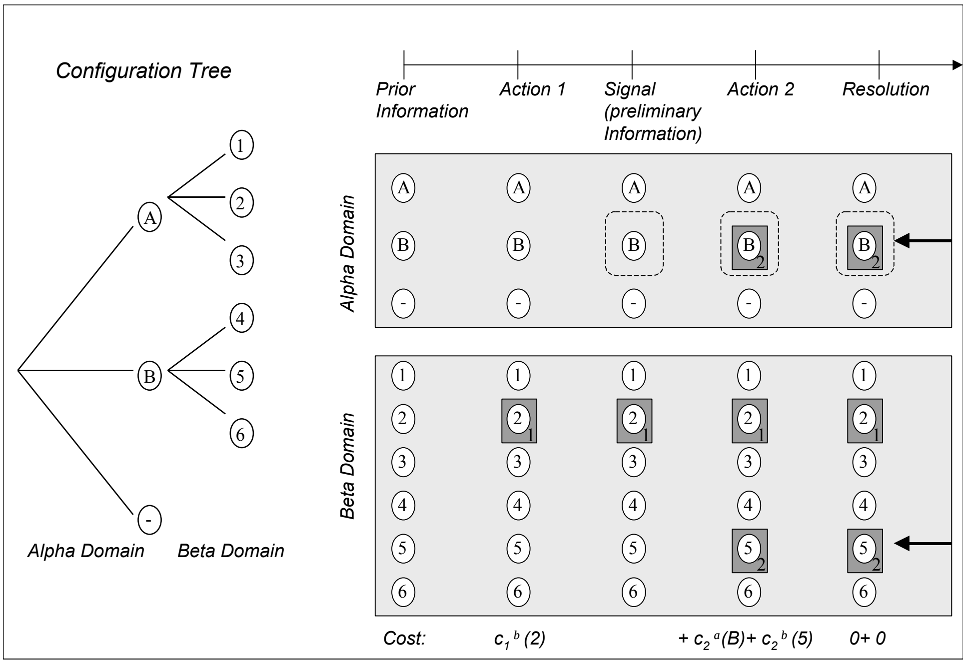

Figure 5: Example of the hierarchical case (Legend see Figure 3)

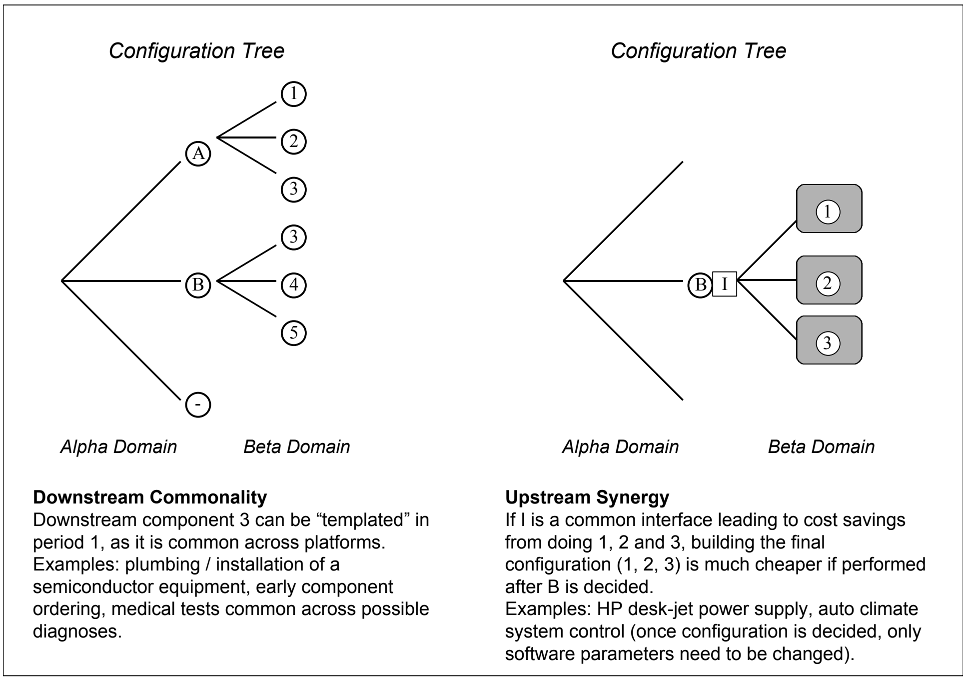

Figure 6: Examples for component commonality and platform synergy