Holding or Folding?

R&D Portfolio Strategy Under Different Information Regimes

Francesca Gino and Gary Pisano[1]

Harvard Business School

April 18, 2005

Abstract

One of the most vexing challenges in the management of R&D concerns decisions whether to continue or halt development of a project. Such decisions must often be made in the face of significant uncertainty regarding the technical or commercial feasibility of projects and aggregate resource constraints. To reduce such uncertainty, firms should exploit the information that becomes available in the development cycle. In this paper, we argue that in different R&D contexts there are significant differences in the timing with which information becomes available during the development cycle. In some contexts, information becomes available relatively quickly due to advanced prototyping and test technologies and the availability of powerful predictive models. In others, information becomes available relatively late in the R&D process. We use the term "*information regime*" to describe these different inter-temporal patterns of information availability over the course of the development cycle. Our thesis is that R&D performance is influenced by the match between risk management heuristics and the information regime in which the firm operates. We test this argument using a simulation model of the R&D process.

[1] Financial support from the Harvard Business School Division of Faculty Research and Development is gratefully acknowledged.

I. Introduction

The management of R&D project portfolios represents an important and vexing challenge for many organizations. Wheelwright and Clark (1992), for instance, suggest that poor management of R&D capacity utilization can seriously damage R&D performance. They also report from their field based research that poor planning, lack of discipline in project screening, and inadequate attention to strategy is a fairly common problem. R&D project portfolio management is difficult because decisions about individual projects must often be made in the face of significant uncertainty regarding the technical or commercial feasibility. In addition, because of aggregate resource constraints, these decisions are not independent across projects. Adler et al (1995) show that when projects share resources "congestion effects" can lead to significant delays in project completion times and output rates. To deal with these problems, organizations typically divide R&D projects into discrete phases, the end of each providing an opportunity to evaluate information on progress and to decide whether to continue or terminate the project. And most companies utilize some type of "portfolio review" processes to allocate resources across projects.

An important strategy variable in these processes concerns the criteria for advancing or terminating projects at different development stages, and specifically, the amount of risk the organization is willing to take. Because all R&D projects involve uncertainty to varying degrees, decisions about whether to terminate a project or advance it to the next stage of development may come down to different risk strategies. The pharmaceutical industry, for instance, is full of examples of companies coming to very different conclusions about the technical and commercial prospects of very similar drugs. When faced with similar data, one organization might choose to continue development while another might choose to terminate the project. Despite the development of increasingly sophisticated analytical tools for project selection and resource allocation (e.g. real option valuation), managerial judgment continues to play a pivotal role in these decisions. Such judgments are generally expressed through the heuristics --or simplifying strategy or "rule of thumb"-- deployed in the organization (Cyert and March (1963), Simon (1957), Nelson & Winter (1977)).

This paper examines the performance implications of risk management heuristics influencing decisions to advance or terminate projects at various phases of the development process. The strategies embedded in these heuristics span a continuum. At one extreme, there are strategies focused on weeding out questionable projects as early as possible, and committing only to those select few projects that appear highly likely to succeed. Such a strategy sets a "high bar" for continuing development on any given project. Because this strategy tends to result in significantly more projects being terminated early in the development cycle, we refer to it as "fast kill." At the other end of the spectrum are "slow kill" strategies which impose a lower threshold for continued development. Under the "slow kill" strategy, projects with less certain prospects are allowed to continue development until later phases of the process. "Tough" decisions are made late in the development cycle. The obvious trade-off between "fast kill" and "late kill" strategies concerns the relatively likelihood of "type I" and "type II" errors.

In coming to grips with the appropriate strategy, managers typically consider the relative costs of each type of error. The greater the cost of advancing a project into later phases of development, the more the firm would presumably error on the side of caution (and vice versa). Another factor also needs to be considered: the *timing with which information becomes available* in the development cycle. In this paper, we explore how differences in this latter dimension influence the performance of different R&D risk management strategies. We argue that in different R&D contexts there are significant differences in the timing with which information becomes available during the development cycle. In some contexts, information becomes available relatively quickly due to advanced prototyping and test technologies and the availability of powerful predictive models. In others, information becomes available relatively late in the R&D process. We use the term "*information regime*" to describe these different inter-temporal patterns of information availability over the course of the development cycle.[2]

[2] Our concept of information regime is similar to Krishnan et al's (1997) notion of "degree evolution." However, their concept focuses on uncertainty about specific design parameters of an evolving product. Our concept of information regime deals with information regarding the prospects of an entire project rather than specific design

Our thesis is that R&D performance is influenced by the match between risk management heuristics and the information regime in which the firm operates. We test this argument using a simulation model of the R&D process. We then discuss implications for future research on R&D portfolio and offer some tentative suggestions for managerial practice.

While previous studies have explored the impact on information availability on project management decisions such as the degree of overlap between development stages (see, for instance, Krishnan et al (1997), Terwiesch and Loch (1999)), the concept has not been applied to R&D portfolio management decisions. Quite surprisingly, the problem of managing R&D portfolios has received much less attention than portfolio selection. A common assumption in the studies addressing this issue is that a given set of projects is initially specified (see, for instance, the study by Hopp (1987), Banerjee and Hopp (2001), Subramanian et al. (2000)). Our model looks at the problem of portfolio management but, different from previous studies, it allows the set of initiated projects to vary over time. In this paper, we take a behavioral perspective and assume that the shear complexity of R&D portfolio decision-making forces precludes optimization in any meaningful sense. Thus, our focus is on understanding the performance of different heuristics rather than trying to derive an optimal strategy. Our paper departs from previous work on R&D portfolio also in another respect. Prior work has explored the optimal number of parallel approaches to solving a specific design or technical problem (Nelson (1961), Marschak et al (1967), Ding and Eliashberg (2002)). This work differs from our own in that we are interested in the management of a portfolio of independent projects, whereas the previous work focuses on "projects" geared toward the same end.

---

parameters. This distinction is appropriate given our different focal problems. Krishnan et al. (1997) were concerned with the problem of the optimal overlap between stages of the development process. In our model, we assume a sequential process, and are concerned with whether or not a project is advanced from one stage to the next.

II. Conceptual Framework: R&D as Information Processing

All R&D is, by definition, uncertain. Uncertainty in R&D can take a variety of forms. For instance, in drugs, the vast majority of projects end in failure at different phases of the development process because the molecule under development does not meet safety or efficacy goals. In other contexts, like automobiles or electronics, product engineers face uncertainty about the feasibility and relative technical and commercial attractiveness of various design options. There can also be uncertainty about user (customer) preferences, what competitors are doing, and other elements of the external environment. One of the fundamental challenges of R&D management is *sorting* projects or specific design solutions according to their attractiveness (i.e. selecting the winners and killing the losers) under uncertainty.

In R&D, sorting occurs through a series of problem-solving/experimental cycle. These cycles utilize various methods such as computer simulations, laboratory experiments, prototype tests, and market research. Regardless of the method, these are all modes of *generating information*. We can think of each problem-solving cycle as "adding" information to the product, or, more formally, narrowing the distribution of expected outcomes of different options (Krishan et al (1997)). In essence, each round of experimentation/testing generates data. These data are then transformed into information via a process of interpretation. We define information in entropic terms, following the literature on the theory of communications and information (Shannon and Weaver (1949)): information is equivalent to a reduction in uncertainty and it can be measured by entropy. From this perspective, R&D is an information processing process (e.g. Clark and Fujimoto (1991), Krishnan et al (1997), Terwiesch and Loch (1999)).

The following example should help to illustrate. Consider a firm that is attempting to develop a drug to treat migraine headaches. The firm has discovered a molecule which, based on information in prior literature, it believes will safely and effectively treat the condition, and be successful in the market. At the beginning point of the project, we can think of the firm as having relatively little information about the ultimate prospects of success of the drug. Based on prior experience, it might know, for instance, that the probability of a molecule at this stage ever reaching the market is one in five thousand. But, it has

little information about whether this is the one in five thousand. The drug could be a wonderful success, but it could also be a dismal failure. Posed formally, the probability distribution of potential future outcomes is wide (high variance).

The firm updates its information on the project by investing in various kinds of testing (e.g., first tests in laboratory animals and later in humans). Each round of tests or experiments generates raw data about the drug's safety and efficacy. These data present a distribution of outcomes (e.g. the percentage of patients who responded favorably to the drug over various time intervals). These data need to be interpreted to update the firm's expectations about the prospects of the drug. To the extent these data allow the firm to narrow down the probability distribution of different outcomes for the drug (i.e. the chance of it being a "winner" or "loser"), they have provided useful information. They have reduced uncertainty.[3] The challenge is to decide, at various points in the development process, whether it is worth continuing or abandoning investment in the project. Note, reducing uncertainty does not necessarily imply improving the odds of success. In fact, it can mean quite the contrary. A study of how the drug works in 100 patients, for instance, that showed that only 2 people responded at all to drug, and 98 had no response, would dramatically reduce the uncertainty of the project; it would likely make the decision makers quite certain that they had a "loser" on their hands!

We can think of this cycle being repeated throughout the development cycle. With each round, the firm updates its information, and the probability distribution narrows. And, it becomes more confident in the likelihood of success (or failure). While seemingly stylized, this depiction is quite reflective of the specific attributes of the process in industries like pharmaceuticals where firms systematically collect data throughout the project. However, we also think it reflects the essential

[3] In information theory, the concept of information has evolved since Shannon's work so that a more general definition has been identified: information is a collection of data altering a representational construct (Daft and Macintosh (1981), MacKay (1969)). In other words, the effect of information is a change in a mental representation. This implies that data and information are two different concepts: In particular, data become information when they are given a meaning or interpretation for activities to be undertaken (Daft and Macintosh (1981)).

attributes of R&D processes in a wide range of contexts where uncertainty regarding project viability or specific design attributes is resolved over a series of experiments or other data acquisition activities (e.g. market research).

The information based conception highlights two salient characteristics of the R&D process. First, there is almost never perfect information about a project. Even as firms accumulate data over multiple rounds of experimentation, they simply narrow down the probability distribution of expected outcomes, but there is usually some "residual" uncertainty. The way information accumulates over the development cycle, however, may vary across contexts. Second, decisions to advance or terminate a project are based partly on available information, but also partly on the decision-maker's risk management strategies. We discuss each of these below.

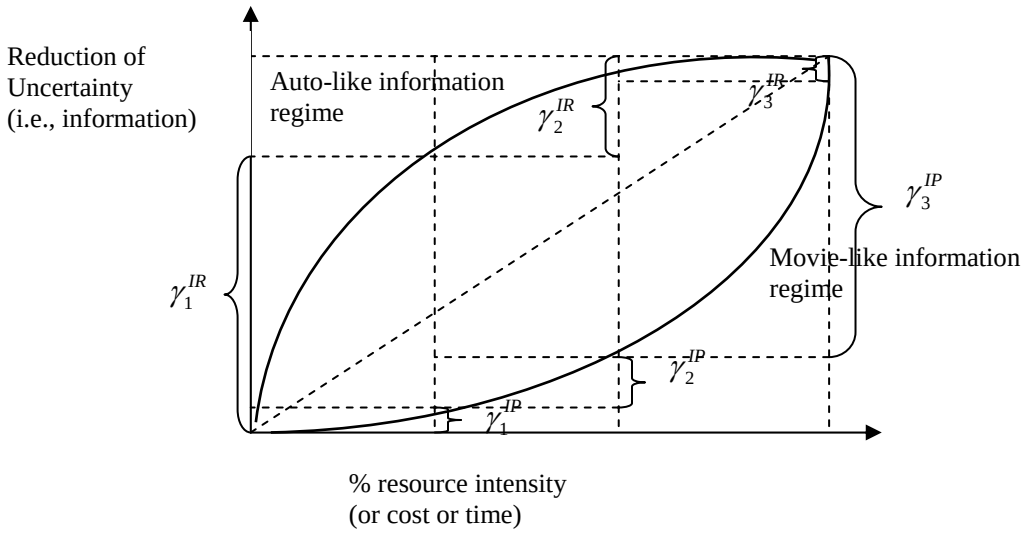
II.1 Information Regimes

The concept of R&D as an information processing process is well established in the literature (e.g. Clark and Fujimoto (1991), Krishnan et al. (1997), Terwiesch and Loch (1999)). This perspective highlights the link between investments in experimentation/testing and uncertainty reduction. In general, we would expect a positive relationship. Experiments create data which create information which in turn reduces uncertainty. But, the linkages between experimentation, data, and information are complex, and are likely to vary substantially across contexts. Let us define the "productivity" of an experiment as the amount of uncertainty reduction per unit of effort of experiment (e.g. hours, dollars, etc.). The productivity of experiments varies across contexts (Thomke (2003)). Consider two contrasting examples: automobiles and movies. At the outset of an automobile development project, there may be uncertainty surrounding the safety performance of various body and chassis designs, but thanks to advancements in simulation technology, it is possible to resolve a significant fraction of this technical uncertainty relatively early in the development process (Thomke and Fujimoto (2000)). In contrast, consider the process of producing a movie. Initially, producers attempt to sort "winners" from "losers" by reading scripts, but the script is a poor predictor of the success of the movie. Indeed, the vast majority of Hollywood films are

failures (DeVany (2004)). Even as the movie begins to "take shape" (the creation of a cast, the filming, the editing, the music score, etc.), it is extremely difficult to predict how well a film will do, until it opens - and even then, there is evidence that the opening is a poor predictor of ultimate success (DeVany (2004)). Unlike automobiles, uncertainty remains high throughout the entire development process, and is not resolved until the very end. Figure 1 provides a stylized comparison of the rate of uncertainty reduction in a context like automobiles versus one like movies.

The relationship between experimentation effort and the rate of decline in uncertainty over the development cycle constitutes an *information regime.* Any given development project takes place within the context of a specific information regime. In essence, an information regime defines an exogenous set of constraints on the organization's ability to reduce uncertainty via experimentation. We use the term "*information rich*" (i.e., IR) to describe regimes where the marginal rate of uncertainty reduction is high relatively early in the project, and "*information poor*" (i.e., IP) to describe regimes where the marginal rate of uncertainty reduction is low in the earlier phases of the project.

Figure 1. Information regimes.



As shown in Figure 1, the marginal rate of uncertainty reduction is captured by the parameter γ_i^j, where $i = \{1,2,3\}$ and $j = \{IP, IR\}$. Such parameter represents the signal a firm receives about the nature of projects under development, i.e. a signal on whether the project will succeed or fail. An information

rich regime generates a strong signal early in the development process (i.e., γ_1^{IR}) and a weak signal in the latest phase (i.e., γ_3^{IR}). An information poor regime, instead, generates a weak signal in the earliest phase (i.e., γ_1^{IP}) and a strong signal later on (i.e., γ_3^{IP}). That is, assuming that the R&D process can be divided in 3 phases[4] and that the strength of the signal in phase II does not differ for the two regimes, the following relationships hold: $\gamma_1^{IR} > \gamma_1^{IP}$ and $\gamma_3^{IR} < \gamma_3^{IP}$. The strength of the signal provides information on whether the project is really a winner and, consequently, it reduces the uncertainty the firm is facing in the development process. As we discuss later, how the availability of such information impacts R&D performance depends on the firm's risk management strategy, and in particular, the criteria it uses to move projects forward.

Information regimes vary across industries, firms, and even across projects executed within a firm. Terwiesch and Loch (1999), for instance, found the rate of uncertainty resolution to vary significantly across their sample of 140 electronics product development projects. A number of factors are likely to influence differences in information regimes across contexts. One of these is differences in experimentation technology. The amount of information generated by an experiment depends on two characteristics of the experimentation technology, namely the cycle time of the experiment and the degree to which process is capable of generating signal versus noise. Advances in experimentation technology, such as computer-aided simulation, have dramatically reduced the cycle time required for an experimental cycle. Similarly, experimentation technologies that enable more precise manipulations of either prototypes or experimental conditions are generally capable of generating superior signal-to-noise ratios. Experimentation technologies vary dramatically across contexts. For instance, in aircraft, automobile, and semiconductors, sophisticated computer simulation models enable engineers to conduct much of their experimentation virtually. But, at this point in time, such simulation tools are barely in their infancy in

[4] In the present work, we model R&D as a 3-stage process. Such a process characterizes most of the modern companies if we think of phase I as research, of phase II as development, and of phase III as commercialization.

pharmaceuticals. And, it would be hard to imagine how one might conduct a computer simulation of a (non-animated) movie.

A second, and related factor, influencing information regimes is the depth of causal knowledge. An experiment generates data, but to make meaningful predictions from data, one needs a good model that maps experimental observations into predictions. Good models (i.e. highly predictive) require deep knowledge about the underlying cause-effect relationships, and the relevant interactions between variables. Such knowledge can be based on empirical experience (and rules of thumb) or have been captured in scientific theories. The depth of such knowledge can clearly vary across industries. Boeing can use simulation because it has accumulated deep knowledge about how various design parameters map into to airframe performance. This knowledge is partially based on its own extensive experience as well as well established principles of aerodynamics. Drugs in contrast generally represent a realm where such causal knowledge has historically been quite poor, and, as a result, it was difficult to predict how well a drug would perform in humans with extensive testing of the drug in humans. Of course, even within an industry, information regimes can vary dramatically. In drugs, some diseases are much better understood than others. The biological mechanism of high cholesterol is known at the molecular level; the biology of most psychiatric diseases is relatively poorly understood. As a result, the development of medicines for psychiatric conditions is still very much a trial-and-error process.

The role each of these features plays in an information regime can be explored by looking at their manifestation in the case of a strong versus weak signal (see Table 1). Information rich regimes typically utilize advanced experimentation/simulation tools and have well developed "models" that enable early reliable predictions of product performance. Note, in our concept of information rich regimes, later stages of the development process introduce more noise into the process. At first glance, this may seem like an odd assumption. It suggests that one may know less later in the process. In fact, this feature characterizes many development contexts. For instance, in chemical manufacturing, the best time to identify a problem with the basic chemistry of the process is in the laboratory. Laboratory experiments provide a highly controlled, low noise environment in which to evaluate the performance (e.g. yields) of alternative

processes. Once the process is transferred to the plant, a host of additional factors such as human factors, equipment differences, and variance in raw materials create noise which makes it harder to identify the root cause of problems. Similarly, in automobiles, computer simulated crashes actually provide higher quality (lower noise) information than tests of physical prototypes (Thomke (2003)). Thus, our concept of information "richness" is relative: it highlights differences in the signal-to-noise ratio between early and later stages of development.

Table 1. Information regimes features and signal strength.

	Signal Strength	
	Strong	**Weak**
Experimentation Technology	- Short cycle time for experiments - High signal-to-noise ratio in the data generated through exp.	- Long cycle time for experiments - High signal-to-noise ratio in the data generated through exp.
State of Knowledge	Good model for interpreting raw data	Poor model for interpreting raw data

II.2 R&D Portfolio Management Heuristics

The primary challenge in R&D portfolio management is to decide which projects to start (project selection) and which ones to continue (terminate) in the face of new information, taking into account capacity constraints. These decisions involve tradeoffs between risks, returns, and time horizons for payoffs. This process is complicated not only by the uncertainty, but also by the impact of capacity constraints. A decision to advance project "A" to the next stage of development may mean that there are not enough resources to start project "B". In theory, such tradeoffs can be explored with a technique such as dynamic programming. In reality, the sheer complexity, ambiguity, and uncertainty of most companies' R&D portfolios make this an essentially impossible optimization problem for an organization to solve. Information about technical feasibility or market potential is often highly subjective or incomplete. In addition, the time and resources required to complete a given stage may also be far from certain. And then, we need to consider that for a large firm, these issues must be considered simultaneously for a portfolio of many projects. While the literature has proposed various decision-

theoretic models, they have not become commonly used in practice because of their complexity (Loch and Kavadias (2002)).

As is well known in the behavioral decision-making literature, organizations and individuals rely on heuristic when confronted with highly complex problems (Simon (1957), Cyert and March (1963), Bazerman (2005), Kahneman & Tversky (1974)). R&D portfolio management is no different. Case studies of R&D portfolio illustrate this point (Fleming, Pisano, and Strick (2004)). Our interest in this paper lies in the heuristics that influence decisions whether to continue or terminate projects at different stages of the development process.

For simplicity, let us assume that projects fall into two categories: "winners" and "losers" where a winner achieves some target level of commercial and/or technical performance. Going back to our earlier discussion, the purpose of R&D is to generate information that sheds light on whether any given project is a winner or loser. However, because of incomplete information, what a firm observes is a distribution of potential outcomes based on the data. The challenge then is to decide, in the face of such uncertainty, whether to move the project forward or to terminate it. We can think of a portfolio heuristic as a figurative "bar" for sorting projects. Projects "below" the bar are terminated, while those "above" the bar continue to receive investment.

More formally, in statistical terms, the threshold for continuation is akin to the confidence interval in hypothesis testing. We might think of the null hypothesis as taking the form: "the project is a winner" (where, for simplicity, we assume the firm has a set of criteria for defining a winner: e.g. sales growth, profitability, technical performance, etc.). At any given point in the project, depending on the available information, there is a distribution of potential outcomes for the project. Of course, there is uncertainty. The manager's challenge is to decide whether to accept or reject the hypothesis that this particular project is a winner (or loser) based on the available information. Here, the manager's problem is very similar to that of a researcher confronted with a set of empirical or experimental data. The decision whether to accept or reject the hypothesis depends partially on the choice of confidential interval or critical value. A "low bar" expands the probability that the hypothesis will be accepted, but this also

increases the chance for a type II error (i.e. accepting a hypothesis that is false). The same is true in a development project. If the firm sets a "low bar", it is essentially taking a bigger risk of advancing a project that is really a "loser." The converse is also true. A firm that sets a "high bar" is selecting projects with relatively high probabilities of success. It reduces its chances of advancing a loser, but also increases its chance of inadvertently killing a project that was really a winner. The manager faces an additional challenge; there is an opportunity of accepting the hypothesis. A decision to advance a project consumes resources which are then not available to other projects.

Managers are often exhorted to be more rigorous in their selection of projects early in the development cycle. These are sometimes referred to as "fast kill" strategies because they focus on killing projects early in the development cycle. The logic behind a "fast kill" strategy is compelling. As projects advance through the development process they consume a greater level of resources. Sticking with a loser is costly, and given capacity constraints, may prevent the firm from starting other, more attractive opportunities. However, there are trade-offs. Fast kill is also likely to lead to the premature termination of projects that would have resulted in a winner. As we argue below, the performance of any given heuristic is contingent on the information regime.

III. A Contingent Model of R&D Portfolio Strategy

The firm's challenge is to maximize its R&D output, subject to an R&D budget constraint (B)[5]. We assume the firm operates a multi-stage R&D process (we will use 3). The firm's challenge is to decide which projects to advance from one stage to the next, and which ones to terminate. Each stage of the R&D process is associated with an expected cost (c_j). We assume that $c_3 > c_2 > c_1$. That is, the least expensive stage is first, the most expensive stage is last. This assumption corresponds reasonably

[5] $B = c_1x_1 + c_2x_2 + c_3x_3$ where x_j indicates the number of projects in Phase j and c_j indicates the cost of Phase j. At the beginning of each simulation, there are only projects in Phase 1 and thus the budget constraint is equal to $B = c_1x_1 = M$.

well to the cost structure of R&D in most industries. The firm pursues a risk management heuristic (H) which incorporates a vector of probability thresholds for advancing the project at each phase ($H = f([p_1^S, p_2^S, p_3^S])$). For instance, if the firm has a threshold for advancement in Phase 1 (i.e., p_1^S) of .50, any project with a subjective probability assessment above .50 will be advanced; any below that level will be terminated. Thus, the firm's returns π or output depend both on its heuristic H and the nature of the projects it invested in (whether they will succeed or fail). For simplicity, we assume that all successful projects have the same returns. Thus, the firm's objective is to maximize the number of successful product launches given its budget constraint.

We can think of the firm as drawing from a metaphorical "urn" of project candidates. At the outset, a project is either a "winner" or a "loser", however, there firm does not perceive this. As the firm invests in R&D, it acquires some signal about the probabilities that a particular project is either a winner or loser. The signal γ_t conveys information about the true nature of projects at time t. The stronger γ_t, the easier it is for the firm to distinguish between winners and losers. γ_t varies by the information regime in which the firm finds itself. What the firm observes is a subjective probability estimate at each point that the project will eventually be successful. It makes decisions at the end of each phase whether or not to continue the project based on its heuristic. We assume that at the end of the process (the end of the 3rd phase) the "truth" is revealed about any remaining projects, and any "losers" that made it through the process are at that point eliminated. While seemingly a stylized description, it actually fits quite accurately with what happens in an industry like pharmaceuticals where a 3rd party regulator (the FDA) ultimately approves new products. In other contexts, we might think of customers (the market) playing ultimately deciding which products are "winners" and "losers."

There are two components driving performance: overall output of successful projects (successes) and the costs incurred over that period. Overall output of successful projects is determined by two factors. The first is the ex ante probability of drawing a winning project. Metaphorically, we can think of this as the relative number of "winning" project in the urn. In reality, we can think of this as the fertility

or opportunity available in the firm's technological environment. A firm that draws from a more "fertile" environment should, all things being equal, achieve a greater number of successful launches. The second factor influencing output is the firm's accuracy in killing projects. Specifically, its output will be reduced if it inadvertently kills projects that were actually winners. The chances of killing a "winning" project are a function of the firm's selection heuristics and their interaction with the information regime. The number of "killed winners" should increase as the probability threshold for advancement increases. That is, the higher the confidence required for advancement, the more likely the firm is to commit a type II error and kill a winning project. Output also declines if the firm fails to identify and terminate losers early in the development process because these projects consume resources; in essence, failing to kill losers early in the process reduces effective development capacity.

Total costs incurred fall into two categories: total sum costs incurred in developing the successful projects and total cost incurred in developing projects that were eventually terminated. Each successful project incurs a cost of $c_1 + c_2 + c_3$. The cost of each terminated projects depends on the stage at which the project was terminated. The cost of a project terminated in Phase *j* is c_j. The cost c_j increases with *j*, so projects killed later impose greater costs. Note, the costs of killing a project are the same whether the project was really a loser or really a winner, although a killed winner has the added opportunity cost of foregone returns from a successful product.

We are now in position to consider the trade-offs imposed by different sorting heuristics. For simplicity, we will consider heuristics at polar extremes: Late Kill and Fast Kill. A "late kill" strategy results in advancing the project even if its subjective probability for success is relatively low. In practice, this results in relatively few projects being killed earlier, and more projects being killed later, or even not being killed at all (i.e. letting the market decide). The benefit of the late kill heuristic is that it is less likely to kill winners, and this potentially increases output. However, the late kill heuristic also results in a greater likelihood of losing projects making their way into later development process. A project killed after Phase III incurs costs of $c_1 + c_2 + c_3$. Late kills have two deleterious effects on productivity. First,

they increase R&D costs. Second, because of budget constraints, they reduce effective capacity and this reduces output over time. For fast kill heuristics, the firm decreases its chances of wasting resources on losing projects. However, it increases its odds of inadvertently killing a winner. This will result in a reduction in total output.

Risk strategies are also likely to impact the "balance" in the firm's pipeline. With a fast kill strategy, the firm will tend to have a front-loaded pipeline (i.e. relatively more projects in early phases than later phases). The opposite should hold for a late kill strategy. The effect of risk strategies on the "balance" in the firm's pipeline is then either exacerbated or reduced based on the information regime the firm operates in. When information becomes available affects the composition of winners relative to loser in the pipeline by influencing the accuracy of screening. To oversimplify, while risk strategies impact the *quantity* in the screening from phase to phase, information regimes impact the *quality* of such screening. To maximize output, there should be a match between the quantity of projects at each stage and the quality of screening at stage. Posed simply, a firm should have more projects in the stages of the process that generate the most information. The performance of any given heuristic is influenced by its accuracy in sorting winners and losers. Ideally, the firm would want to kill all losers as early as possible, without killing any winners. But, given uncertainty, perfect accuracy is impossible. Errors (both type I and type II) will be inevitable and will be more likely to occur when decisions are based on noise. This suggests that the decision to terminate a project (which is irreversible) is best made when the signal strength is highest. That is, errors can be reduced if the firm makes good use of the signal it gets in each of the phases of the development cycle.

This reasoning led us to formulate the following hypothesis: *firms' risk strategies perform significantly better in terms of R&D output when they match the information regime*. In an information rich regime, the signal strength is highest after the initial phase of development. A fast kill strategy exploits this information, and should thus lead to more accurate sorting decisions (and thus higher output). In contrast, in an information poor regime, the signal strength is highest in the late phase of development. Because a late kill strategy exploits the information that becomes available late in the

process, it should lead to more accurate sorting, and thus perform better in this type of information regime. In order to test such hypothesis, we developed a simulation model, which is described in the next section.

IV. Simulation Model

We model a 3-stage R&D process (see Figure 2).[6] The basic parameters of our R&D process are the number of phases, the cost, lead time and attrition rate for each phase, and a fixed total R&D budget. There is uncertainty around both lead times and cost. To capture such uncertainty, we model the time required to complete a phase as a random variable following a Weibull distribution.[7] At each point in time (in the simulation, time is expressed in terms of months), the firm sustains a cost equal to the average monthly cost for that phase. Since the phase lasts for a number of periods equal to the lead time, the overall cost the firm sustains to work on a project in a certain phase is the product between the average monthly cost and the lead time. Settings with different cost and lead time structures can be explored by varying those parameters (see Gino and Pisano (2004)).

Several parameters govern the behavior of a modeled firm: the total time, the portfolio size (which represents the firm's budget constraint), the thresholds a firm uses to move projects forward in the development process, and the signal strength (i.e., the information regime a firm operates in). We have intended this model to be highly general and adaptable to a wide range of particular circumstances. In the simulations we parameterize our model using data from the pharmaceutical industry. The choice of

[6] The model was first developed by Gino and Pisano (2004) to study the impact of firm behavior on R&D output and volatility. In this paper, the model has been adapted so as to explore how different risk strategies perform in different information regimes.

[7] The Weibull distribution is often used to model "time until failure" (for instance in engineering work) because of the many shapes it attains for various values of its parameters. In our simulation (as in Gino and Pisano (2004)), we chose the parameters so that the probability of obtaining a lead time longer than the average one observes in the pharma industry is higher than the probability of obtaining a shorter lead time, for each project and in each phase.

pharmaceuticals was due primarily to the ample published data available on lead times, development costs, and attrition rates for the three clinical trials phases. In particular, on average 17% of the projects are successful (in our terms, thus, the fertility of the environment is equal to 17%). As for cost structure and lead times, we used the values reported in Table 2 below. However, the model is robust enough to be used to explore R&D performance in an extremely wide range of contexts. We chose a simulation approach because of the complexity and analytical intractability of the problem formulated above – especially when we allow the possibility of an infinite number of projects (the process we model is indeed non-Markovian). Through computer simulations we explore the impact of different risk strategies on R&D output, in information poor and information rich regimes.

Table 2. Data used to set the parameters in the simulation.

	Average Values (in each phase)		
	Phase I	**Phase II**	**Phase III**
Lead Time per Phase	21.6 months [st dev = 6.6]	25.7 months [st dev = 11]	30.5 months [st dev = 9.9]
Cost per Phase	MM $ 15.2	MM $ 23.5	MM $ 86.3

* *Source*: J.A. Di Masi et al., "The Price of Innovation: New Estimates of Drug Development Costs", *Journal of Health Care Economics*, 22, 2003.

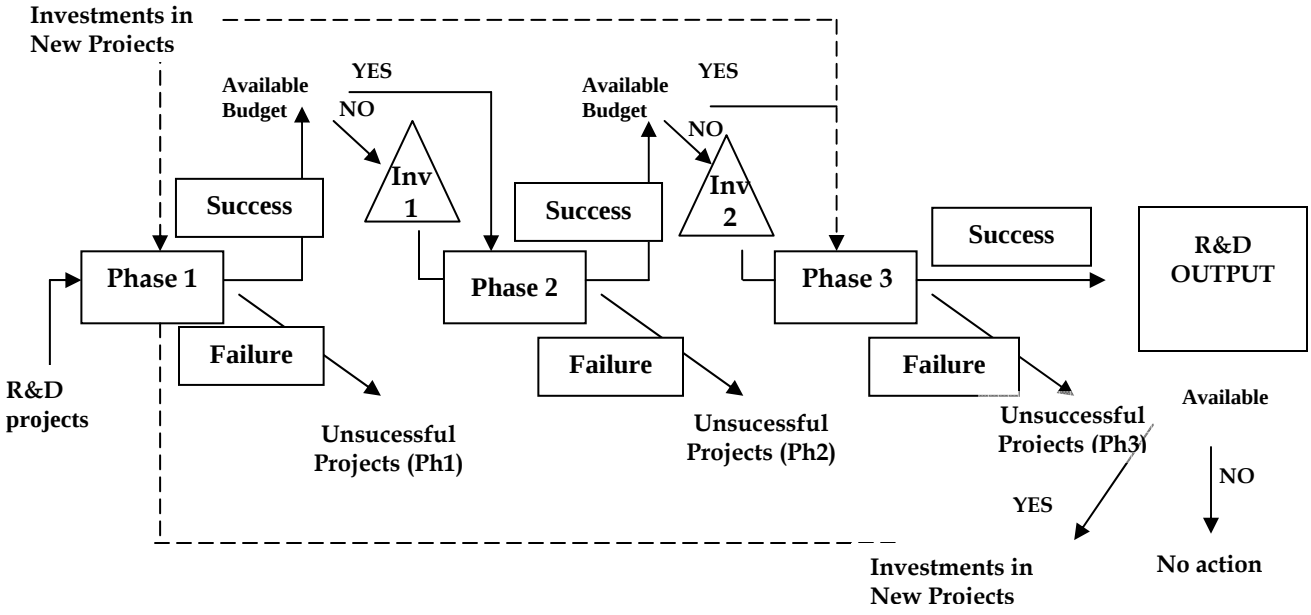
Each simulation set consisted of 100 iterations, in each of which the total time T was set to 1920 months (i.e., 160 years). Results were then averaged over the 100 runs.[8] At the beginning of each simulation, M projects are randomly generated based on the fixed budget. Each project is modeled as a vector, with each element representing a certain feature of the project. In particular, each project is characterized by the following features: a project ID; its cost in each phase; its lead time in each phase

[8] As in Gino and Pisano (2004), we set T equal to 1920 in order to allow a "warm-up period" and thus comment our results after the system has reached a steady state. 160 years seemed to be a reasonable long time to be used as a stopping rule. In pilot simulations we also checked that the use of longer time periods for T did not produce significantly different results.

(i.e., the time required for a project to complete that stage); and an index indicating the current stage a project is in. Projects are independent identical distributed random variables since some of their features are generated randomly. That is, we are implicitly assuming a firm's portfolio is diversified, with each project assuring a fixed return if successful. As soon as projects are generated, they enter the 3-stage process represented in Figure 2.

Projects get funded based on a certain prioritization rule: projects closer to the market get first dibs on resources, and prioritization declines as projects move further back in the "pipeline". That is, the firm always spends resources on the latest phase projects first, the second to latest phase projects second, and so on. Moreover, it can only start a new project if it has resources left over that have no already been committed to phase III, II and I projects. In previous work (Gino and Pisano (2004)), we found this rule optimized output compared to other resource allocation rules. During a simulation run, whenever the firm's budget allows for additional investments, new projects are generated.

Figure 2. A 3-stage R&D process.



In our simulation, when a project is generated, it is labeled randomly as either a "winner" or a "loser". The probability of obtaining a winner in each random draw is equal to a certain percentage number (which we varied across different simulations), which represents the environment fertility. At

each phase, and without knowing if the project is a winner or a loser, the firm must decide whether to advance or kill the project, based on data generated during the phase. Data take on the value of either 0 or 1, to indicate failure and success respectively. If the project was originally a winner, then the firm receives a signal (γ). In particular, the signal was modeled as an increment in the probability of getting a winner in the data generated during the phase. The signal strength varies based on the information regime. Risk strategies are then modeled as *threshold for advancing projects*. Thus, the probability of technical success is used as a measure of risk. We analyze the results of our simulation in the next section.

V. Analysis

How effectively do different risk preferences perform in different information regimes? In this section we explore this question by analyzing the performance of firms' risk strategies in both information poor and information rich regimes. In each simulation run the strategy a certain firm adopts remains the same in each time period. The values used as parameters in each simulation set are summarized in Table 3 and 4. Table 3 indicates the values for the signal strength characterizing each information regime we modeled.

Table 3. Different information regimes.

	Values for the signal strength in each phase		
	Sigma 1	**Sigma 2**	**Sigma 3**
Information Poor Regimes	0.1	0	0.9
	0.2	0	0.8
	0.3	0	0.7
	0.4	0	0.6
	0.5	0	0.5
Information Rich Regimes	0.6	0	0.4
	0.7	0	0.3
	0.8	0	0.2
	0.9	0	0.1

Table 4 presents the thresholds we used to model risk strategies. In our simulations, a firm using a fast kill strategy advances a project from phase I to phase II only if the probability of success is higher than 80%. A firm implementing a late kill strategy, instead, advances a project from phase I to phase II only if the probability of success is higher than 20%.

Table 4. Different types of firm's risk strategies.

	Values for (ex-ante) threshold used in each phase (a project progress only if its probability of success is higher than the threshold, the threshold being the attrition rate)		
	Phase I	**Phase II**	**Phase III**
Late Kill	20 %	50 %	80 %
Fast Kill	80 %	50 %	80 %

We initially simulated the performance of two portfolio sizes: M = 50 and M=10. We chose these values for the approximate correspondence to a prototypical large pharmaceutical company (like Merck) and an emerging biotechnology company. For the ex ante probability of drawing a winner (what we called the "fertility of the environment"), we used p=.17. We chose this value because it corresponds to the reported probability of drug projects to successfully complete all 3 phases of clinical development (see Table 2).

Figure 3 show how different risk strategies perform in different information regimes for the "large firm" portfolio (M=50). As indicated, fast kill strategies outperform late kill strategies in information rich regimes, while late kill strategies perform better in information poor regimes. As information "richness" improves, the advantage of a fast kill strategy increases. By matching its risk strategy with the information regime in which it operates, a firm can increase its R&D output. Table 5 presents tests of statistical significance for the differences in output of the two strategies under each information regime. As shown, the differences are statistically significant for each regime, thus supporting our initial hypothesis.

Table 5. The effect of risk strategies in different information regimes (M=50; p=.17).

	Mean R&D Output		Mann-Whitney test
Information Regime	**Fast Kill**	**Late Kill**	**p-value**
[0.1, 0, 0.9]	0.3245	0.4595	p<.001 (Z=-6.1047)
[0.2, 0, 0.8]	0.4465	0.4865	p=.058 (Z=-1.898)
[0.3, 0, 0.7]	0.534	0.433	p<.001 (Z=-5.071)
[0.4, 0, 0.6]	0.6225	0.3775	p<.001 (Z=-9.789)
[0.5, 0, 0.5]	0.687	0.325	p<.001 (Z=-11.367)
[0.6, 0, 0.4]	0.7595	0.3005	p<.001 (Z=-11.902)
[0.7, 0, 0.3]	0.84	0.2515	p<.001 (Z=-12.184)
[0.8, 0, 0.2]	0.9155	0.197	p<.001 (Z=-12.241)
[0.9, 0, 0.1]	0.8215	0.147	p<.001 (Z=-12.25)

Figure 4 presents results for the "small firm" portfolio (M=10). Here, a somewhat different picture emerges. While information richness improves the performance of fast kill strategies, there are no statistically significant differences in the two strategies for information poor regimes (see Table 6). It appears that for a small firm, a fast kill strategy is never inferior to late kill, and generally dominates for most information regimes.

Table 6. The effect of risk strategies in different information regimes (M=10; p=.17).

	Mean R&D Output		
Information Regime	**Fast Kill**	**Late Kill**	**p-value**
[0.1, 0, 0.9]	0.071	0.0755	n.s.
[0.2, 0, 0.8]	0.086	0.0925	n.s.
[0.3, 0, 0.7]	0.111	0.081	p=.006 (Z=-2.731)
[0.4, 0, 0.6]	0.1175	0.0735	p<.001 (Z=-4.127)
[0.5, 0, 0.5]	0.142	0.0645	p<.001 (Z=-6.199)
[0.6, 0, 0.4]	0.164	0.073	p<.001 (Z=-7.318)
[0.7, 0, 0.3]	0.169	0.0615	p<.001 (Z=-7.361)
[0.8, 0, 0.2]	0.176	0.0485	p<.001 (Z=-9.109)
[0.9, 0, 0.1]	0.1675	0.024	p<.001 (Z=-10.520)

The difference in effect across firm sizes is worth commenting on. For a small firm, a late stage project consumes a disproportionate share of available R&D budget. Moving a project into Phase III entails relatively high risk because it likely means not starting new projects. A failure of a late stage project has a relatively bigger negative impact simply because the firm has fewer overall projects in its portfolio. In contrast, a larger firm benefits from having enough budget to diversify its portfolio across phases. In the next section, we test the robustness of our results by varying the portfolio sizes, the fertility of the environment, and the thresholds embedded in the risk heuristics.

V.1. Robustness of the results

To test robustness, we ran simulations for the following 6 portfolio sizes: M={5, 10, 30, 50, 95, 100}. For each portfolio size, we varied the fertility environment by using the following 5 probability values, p={.05, .25, .50, .75, .95). We thus were able to test the robustness of our results under 30 different experimental conditions. For each of these conditions, we examined the performance of our two strategies (fast kill and late kill) in each of the 9 information regimes shown in Table 3. Finally, to test whether our results were sensitive to the specific thresholds embedded in the risk strategies, we repeated all the above analyses using the values in Table 7. Following our earlier protocol, each experiment involved 100 iterations of the simulation model.

Table 7. Different values used to model firm's risk strategies.

	Values for (ex-ante) threshold used in each phase		
	Phase I	**Phase II**	**Phase III**
Late Kill (Fast Kill)	30 % (70 %)	50 % (50 %)	70 % (30 %)
Late Kill (Fast Kill)	40 % (60 %)	50 % (50 %)	60 % (40 %)

For purposes of length, we can not report results for all of the above conditions (1620 experiments). These results are available from the authors. The results of these experiments are consistent with the general findings reported in the previous section. Information richness improves the

relative performance of fast kill strategies (see Figures 5, 6, and 7). Changing the fertility of the environment affected output in obvious ways (more fertile environment had higher output). However, there was no moderating effect on the relationship between risk strategy and information regime. Changing the thresholds used for the risk heuristics shifted the point at which "fast kill" and "slow kill" strategies crossed over on the information regime axis (see Figure 8).

Perhaps the most interesting finding of the robustness tests concerned the effect of firm size. In the previous section, we reported results for a portfolio of M=10. Here, we went one step further and analyzed a portfolio of M=5. The effect discussed earlier is even more pronounced for a "very small firm." The "kill early" strategy dominates throughout the entire range of information regimes. We find this result interesting in light of previous research on the project portfolio decisions of pharmaceutical firms. In a study of a sample of pharmaceutical R&D project decisions, Guedj and Scharfstein (2004) found significant differences in the rates at which firms terminate projects early in the development cycle. They found that smaller firms with fewer options in their project portfolios had a tendency to let their projects go deeper into the development process before terminating them. This behavior may be driven by financial market pressures to demonstrate progress or principal agent problems. Our results here show that such a behavior for a small firm is unproductive. To maximize R&D output, small firms should be tougher in early stages of development, not more lax.

VI. Conclusions

In this paper, we have explored how the timing of available information during the development process affects the performance of different R&D portfolio risk strategies. The results of the simulation model suggest that the performance of a risk management heuristic is influenced by the extant information regime. In regimes where information becomes available early in the process, "fast kill" strategies are likely to outperform "late kill" strategies (and vice versa).

There are a number of potential implications of this work for both practice and future research. First, it suggests a contingent approach for developing risk management strategies. While managers are

often exhorted to be more disciplined early in the development process, such an approach is most appropriate in information rich regimes. The results also suggest that firms whose project span a range of information regimes may need to pursue multiple risk strategies. One size does not fit all. Second, the results suggest that managers focus more closely on understanding the information regime in which they operate. Bad assumptions about ones information regime could well lead to inappropriate risk strategies. Third, it is clear from all the simulations that information richness improves the potential of R&D performance; that is, fast kill in information rich regime leads to the highest potential output. This suggests that improving information richness in the development process may be a productive strategy for improving overall development performance. Finally, the differences in effect between very small firms compared to large firms have potential implications for entrepreneurial firms. While such firms are often under pressure by investors to show progress by moving projects into later phases of development quickly, our results suggests that this may actually be counterproductive. Very small firms were uniformly better off following a fast kill strategy, independent of information regime.

The paper contains a number of limitations which suggest further avenues for future work. First, while simulation is a valuable tool for exploring alternative strategies, empirical validation is needed. One potential line of empirical research would be to investigate whether differences in R&D portfolio performance across firms or within the same firm is influenced by the diversification of their projects across information regimes. In general, diversification is viewed as healthy. However, based on our simulation results, it may well be the case that diversification across information regimes could detract from performance (assuming firms use a single heuristic). This is an empirically testable hypothesis. Another line of future research should explore the implications of shifting information regimes over time. In this work, we held the regime constant. However, advances in prototyping and experimentation technology are causing information regimes in some industries to shift (i.e. becoming richer). As a result of these advances, it is likely that information will become an even more important resource in R&D. A better understanding of how to manage and exploit information in R&D portfolio will have high pay-offs for management practice.

Figure 3. The Effect of risk strategies in different information regimes.

Parameters: Initial Portfolio Size = 50; Environment Fertility = 17%

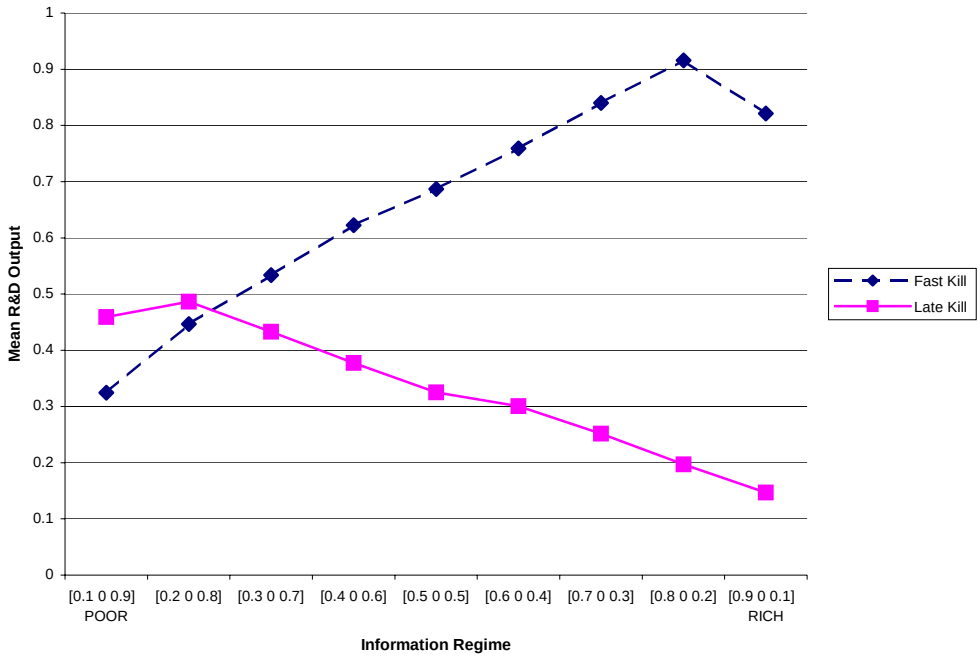


Figure 4. The Effect of risk strategies in different information regimes.

Parameters: Initial Portfolio Size = 10; Environment Fertility = 17%

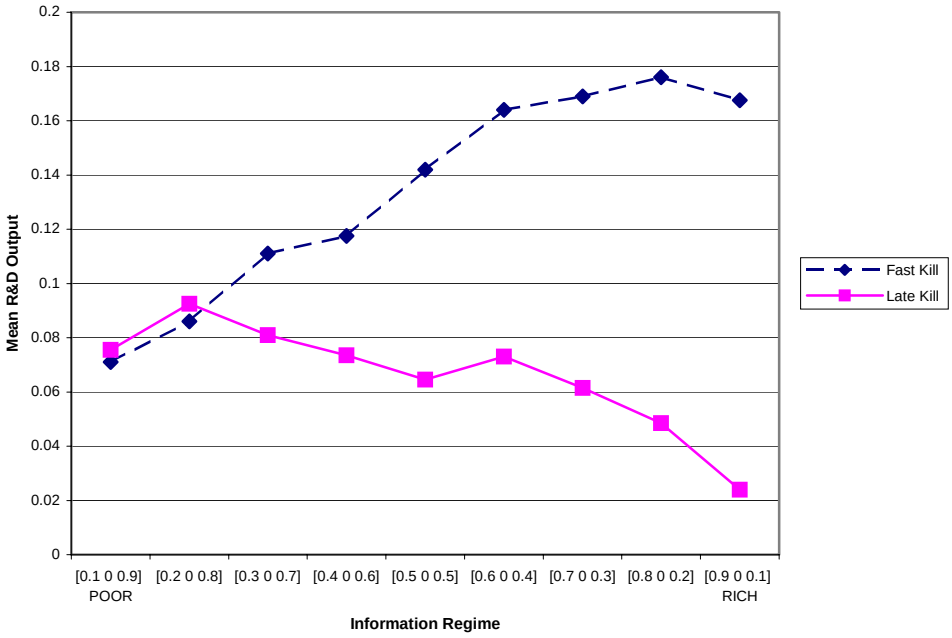
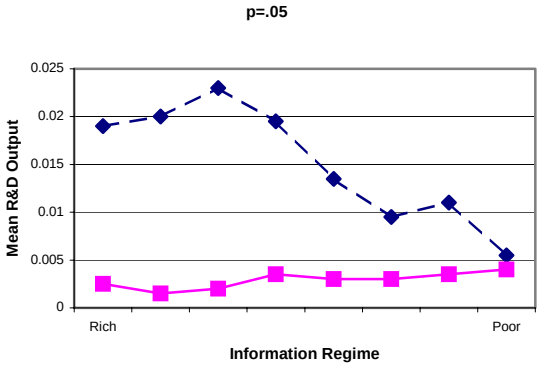
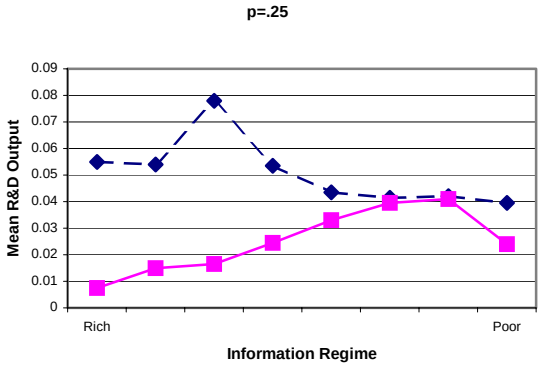
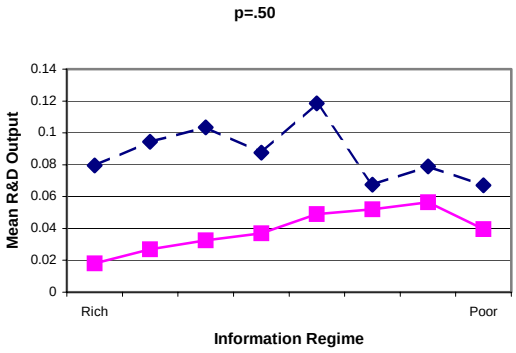
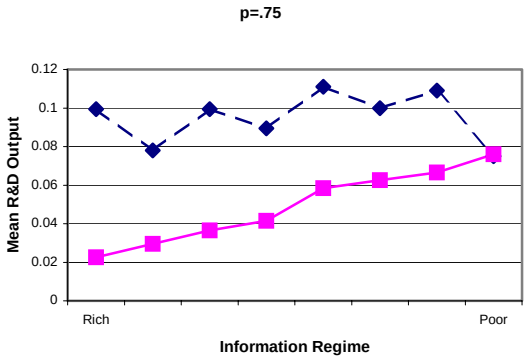


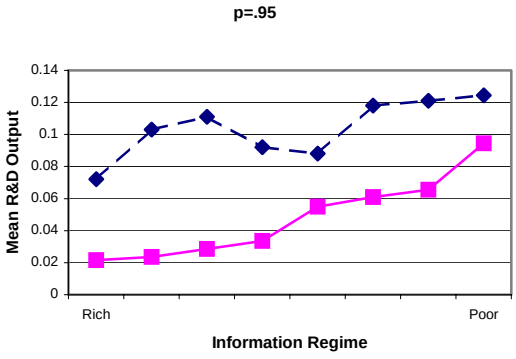
Figure 5. Robustness Tests: Small Portfolio Size (M=5).







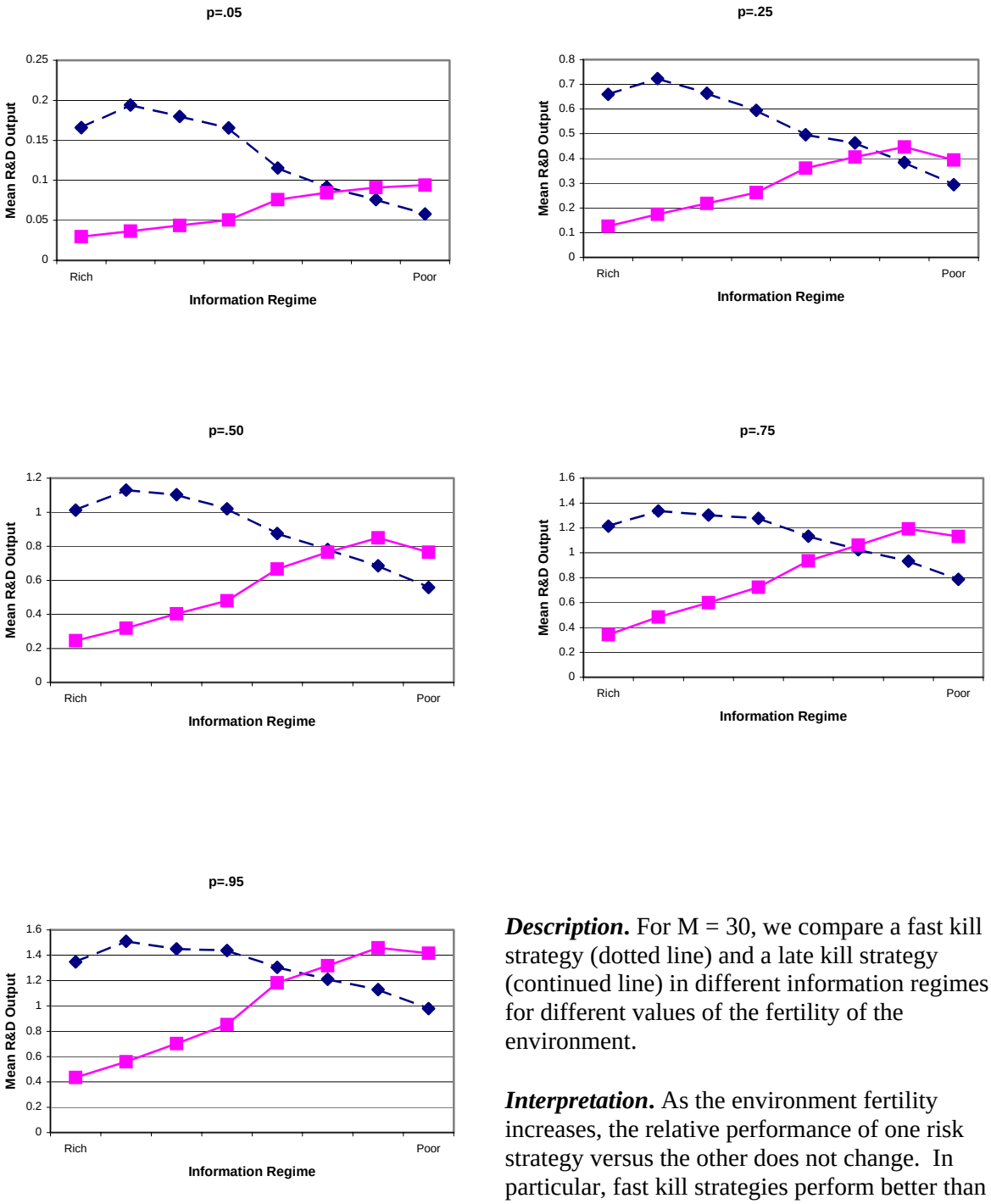




***Description*.** For M = 5, we compare a fast kill strategy (dotted line) and a late kill strategy (continued line) in different information regimes – for different values of the fertility of the environment.

***Interpretation*.** As the environment fertility increases, the relative performance of one risk strategy versus the other does not change. In particular, in the case of small portfolio size, fast kill strategies perform better than late kill strategies both in information rich and in information poor regimes.

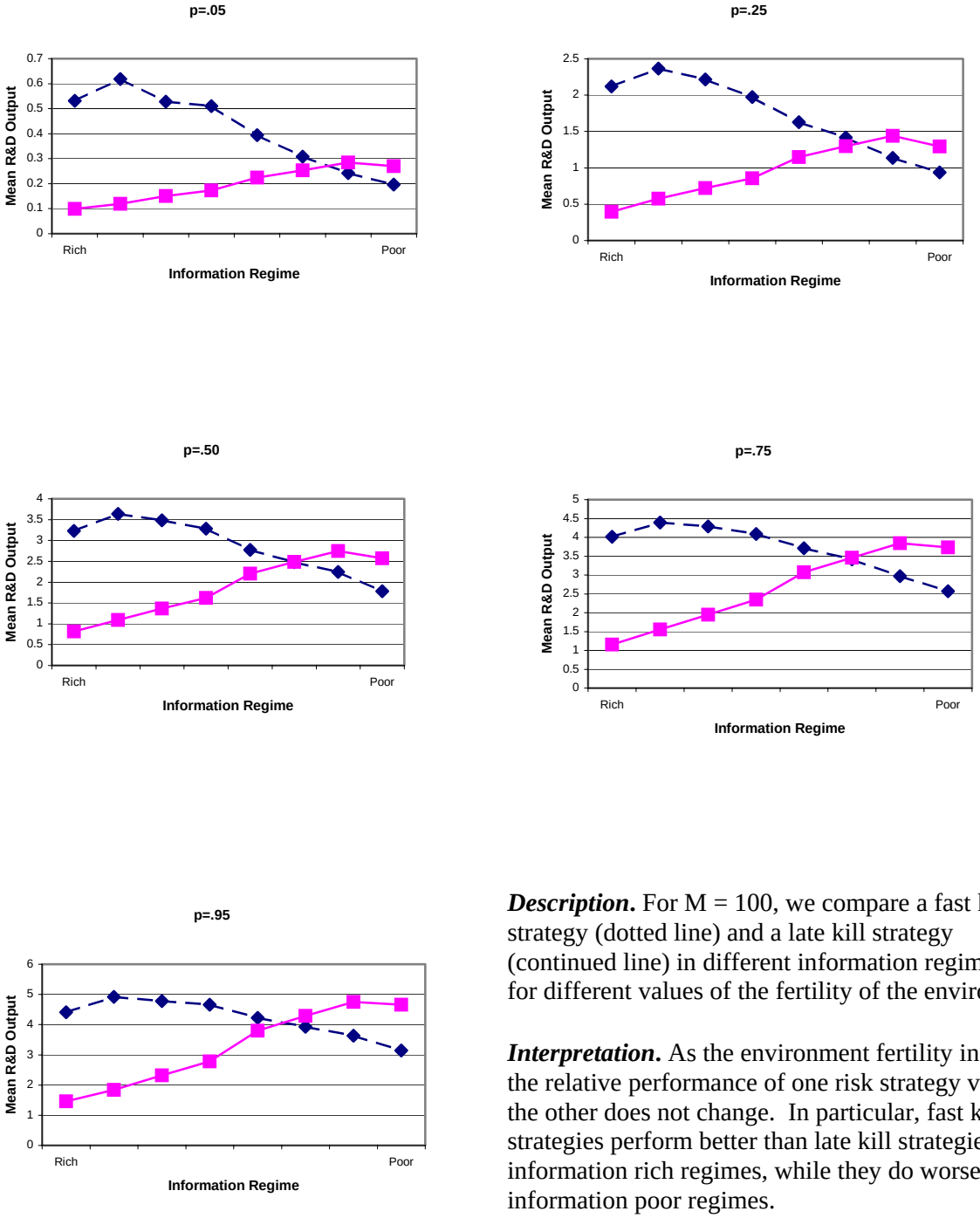
Figure 6. Robustness Tests: Small Portfolio Size (M=30).



***Description*.** For M = 30, we compare a fast kill strategy (dotted line) and a late kill strategy (continued line) in different information regimes – for different values of the fertility of the environment.

***Interpretation*.** As the environment fertility increases, the relative performance of one risk strategy versus the other does not change. In particular, fast kill strategies perform better than late kill strategies in information rich regimes, while they do worse in information poor regimes.

Figure 7. Robustness Tests: Small Portfolio Size (M=100).



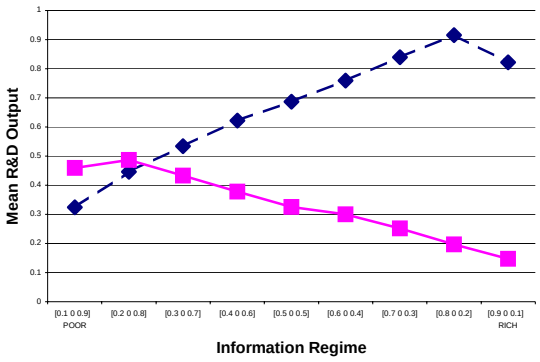
***Description*.** For M = 100, we compare a fast kill strategy (dotted line) and a late kill strategy (continued line) in different information regimes – for different values of the fertility of the environment.

***Interpretation*.** As the environment fertility increases, the relative performance of one risk strategy versus the other does not change. In particular, fast kill strategies perform better than late kill strategies in information rich regimes, while they do worse in information poor regimes.

Figure 8. Robustness Tests: Changing the values for the thresholds embedded in the firm's risk strategies.

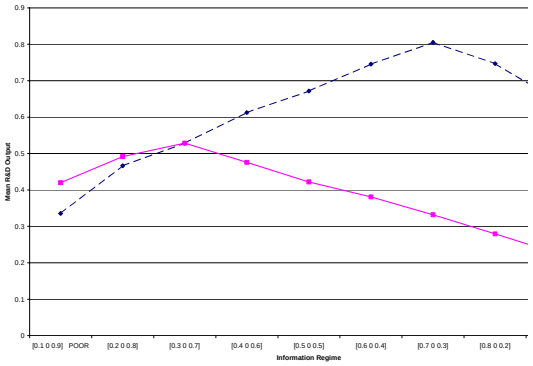
(a) Thresholds:

Fast kill: 80 % (Ph. I), 50 % (Ph. II), 20 % (Ph. III)
Late Kill: 20 % (Ph. I), 50 % (Ph. II), 80 % (Ph. III)



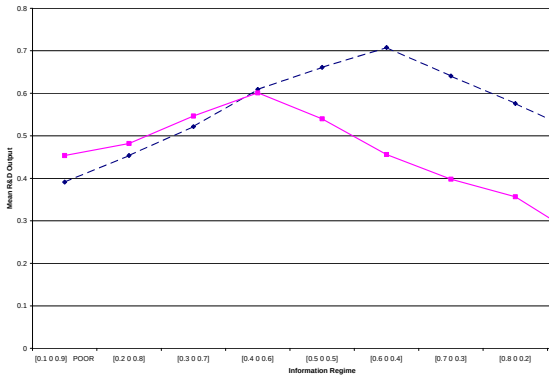
(b) Thresholds:

Fast kill: 70 % (Ph. I), 50 % (Ph. II), 30 % (Ph. III)
Late Kill: 30 % (Ph. I), 50 % (Ph. II), 70 % (Ph. III)



(c) Thresholds:

Fast kill: 60 % (Ph. I), 50 % (Ph. II), 40 % (Ph. III)
Late Kill: 40 % (Ph. I), 50 % (Ph. II), 60 % (Ph. III)



***Description*.** For portfolio size M = 50 and environmental fertility p=.17, we compare different values for the thresholds embedded in the firm's risk strategy in different information regimes. The dotted line represents a fast kill strategy while the continued line represents a late kill strategy.

***Interpretation*.** As the environment fertility increases, the relative performance of one risk strategy versus the other does not change. In particular, fast kill strategies perform better than late kill strategies in information rich regimes, while they do worse in information poor regimes.

References

Adler, P. S., A. Mandelbaum, V. Nguyen, and E. Schwerer (1995). "From Project to Process Management: An Empirically-Based Framework for Analyzing Product Development Time," *Management Science*, 41: 458-484.

Banerjee, S., W.J. Hopp, "The Project Portfolio Management Problem," Department of Industrial Engineering and Management Sciences, 2001, Northwestern University, June.

Bazerman, Max H. (2005). *Judgment in Managerial Decision Making*, John Wiley and Sons.

Clark, Kim and Takahiro Fujimoto. (1991). *Product Development Performance*. Boston: Harvard Business School Press.

Cyert, Richard M., and James G. March, (1963). *A Behavioral Theory of the Firm*, Englewood Cliffs, NJ: Prentice-Hall.

Daft, R.L., and N.B. Macintosh (1981). "A tentative exploration into the amount and equivocality of information processing in organizational work units," *Administrative Science Quarterly*, 26: 207-224.

DeVany, A. (2004). *Hollywood Economics.* Routledge, New York.

Ding, M., and J. Eliashberg (2002). "Structuring the New Product Development Pipeline," *Management Science*, 48: 343-363.

Fleming, Lee, Gary P. Pisano, and Eli Strick (2004). Vertex Pharmaceuticals: R&D Portfolio Management. Harvard Business School Case 604-101.

Hopp, W.J., (1987). "A Sequential Model of R&D Investment Over an Unbounded Time Horizon," *Management Science* 33 (4): 500-508.

Gino, Francesca, and Gary P. Pisano (2004). R&D Performance Volatility: A Behavioral Model. Harvard Business School Working Paper Series, No. 05-015.

Guedj, Ilan, and David Scharfstein (2004). Organizational Form and Investment: Evidence from Drug Development Strategies of Biopharmaceutical Firms. HBS working paper, March.

Krishnan, V., S.D. Eppinger, and D.E. Whitney (1997). "A model-based framework to overlap product development activities," *Management Science* 43 (4): 437-451.

Loch, C. H., and S. Kavadias, (2002). "Dynamic Portfolio Selection of NPD Programs Using Marginal Returns," *Management Science* 48 (10): 1227-1241.

MacKay, D,M. (1969). *Information, mechanism and meaning*, Cambridge, Massachusetts: MIT Press.

Marschak, T., T.K. J. Glennan, R. Summer (1967). *Strategy for R&D: Studies in the Microeconomics of Development*. Springer-Verlag, New York.

Nelson, R. (1961). "Uncertainty, learning, and the economics of parallel research and development efforts." *Rev. Econom. Statist.* 43 (4): 351-364.

Nelson, R., S. Winter, (1977). "In Search of Useful Theory of Innovation." *Research Policy*.

Payne, John W., James R. Bettman, Eric J. Johnson, (1993). *The Adaptive Decision Maker*, Cambridge University Press, Cambridge.

Pisano, Gary P. (2000). Discovering the Future: R&D Strategy at Merck. Harvard Business School Case 601-086.

Rivkin, Jan W., (2000). "Imitation of Complex Strategies", *Management Science*, 46 (6), June: 824-844.

Shannon, C.E., & Weaver, W., (1949). *The mathematical theory of communication*, Urbana: University of Illinois Press.

Simon, Herbert A. (1957). *Models of man - social and rational*. New York, John Wiley and Sons.

Simon, Carl P., and Lawrence Blume, (1994). *Mathematics for Economists.* New York: Norton and Company.

Subramanian, D., Pekny, JF and Reklaitis, GV, (2000). "A Simulation-Optimization Framework for Addressing Combinatorial and stochastic aspects of an R&D pipeline management problem," *Computer and Technical Engineering*, 24, July: 1005 - 1011.

Thomke, Stefan. (2003). *Experimentation Matters: Unlocking the Potential of New Technologies for Innovation*. Boston, Massachusetts: Harvard Business School Press.

Thomke, S., and T. Fujimoto. (2000). “The Effect of 'Front-Loading' Problem-Solving on Product Development Performance.” *Journal of Product Innovation Management* (March).

Terwiesch, C., and C. H. Loch (1999). “Measuring the Effectiveness of Overlapping Development Activities,” *Management Science*, 45: 455-465.

Tversky, A., D. Kahneman, (1974). Judgement under uncertainty: Heuristics and biases. *Science*. 185: 1124-1130.

Wheelwright S C, Clark K B, (1992). *Revolutionizing product development: quantum leaps in speed efficiency and quality*, Free Press, New York.