Research, Development, and Engineering Metrics

John R. Hauser
Sloan School of Management, Massachusetts Institute of Technology, Cambridge, Massachusetts 02142

We seek to understand how the use of Research, Development, and Engineering (R,D&E) metrics can lead to more effective management of R,D&E. This paper combines qualitative and quantitative research to understand and improve the use of R,D&E metrics. Our research begins with interviews of 43 representative Chief Technical Officers, Chief Executive Offices, and researchers at 10 research-intensive international organizations. These interviews, and an extensive review of the literature, provide qualitative insights. Formal mathematical models attempt to explore these qualitative insights based on more general principles.

Our research suggests that metrics-based evaluation and management vary according to the characteristics of the R,D&E activity. For applied projects, we find that project selection can be based on market-outcome metrics when firms use central subsidies to account for short-termism, risk aversion, and scope. With an efficient form of subsidies known as ''tin-cupping,'' the business units have the incentives to choose the projects that are in the firm's best long-term interests. For core-technological development, longer time delays and more risky programs imply that popular R,D&E effectiveness metrics lead researchers to select programs that are not in the firm's long-term interest. Our analyses suggest that firms moderate such market-outcome metrics by placing a larger weight on metrics that attempt to measure research effort more directly. These metrics include standard measures such as publications, citations, patents, citations to patents, and peer review. For basic research, the issues shift to getting the right people and encouraging a breadth of ideas. Unfortunately, metrics that identify the ''best people'' based on research success lead directly to ''not-invented-here'' behaviors. Such behaviors result in research empires that are larger than necessary, but lead to fewer ideas. We suggest that firms use ''research tourism'' metrics, which encourage researchers to take advantage of research spillovers from universities, other industries, and, even, competitors.
(*Marketing*; *Research and Development*; *Product Development*; *Incentives*)

R&D expenditure is often a convenient target when it comes to maintaining or increasing the company dividend. If fact, with R&D expenditure roughly the same amount as the dividend in many companies, it is a significant temptation.

James W. Tipping (1993, p. 13)
Director of Research and Technology, ICI Americas, Inc.

Pioneering research is closely connected to the company's most pressing business problems. . . . Research must ''coproduce'' new technologies and work practices by developing with partners throughout the organization a shared understanding of why these innovations are important.

John Seely Brown (1991, pp. 103–104)
Director of Xerox Palo Alto Research Center (PARC)

0025-1909/98/4412/1670$05.00

Table 1 Managers Interviewed (A total of 43 managers and researchers were interviewed. This table lists some of the titles.)

Organization	Managers Interviewed
Chevron Petroleum Technology	President, Head of Strategic Research, R&D Portfolio Manager
Hoechst Celanese ATG	President, VP Technology, VP Commercial Development, VP Technology & Business Assessment, Director Innovations
AT&T Bell Laboratories	VP Administrative Systems, Director of R&D Programs, Director of Information Applications Architecture
Bosch GmbH	Senior VP for Strategic Planning, Head of Corporate Research
Schlumberger Measurement & Systems	VP Director of R&D, Director of Engineering Process Development, Director of European Tech. Cooperation
Electricite de France	Associate Director R&D, Director of Division
Cable & Wireless plc	Federal Development Director, Director of Technology (HK), Group Strategic Development Advisor
Polaroid Corporation	CEO, Director of Research
US Army Missile RDEC and Army Research Laboratory	Associate Director for Science and Technology, Associate Director for Systems, Deputy Assistant Secretary for Research and Technology/Chief Scientist
Varian Vacuum Products	VP, General Manager

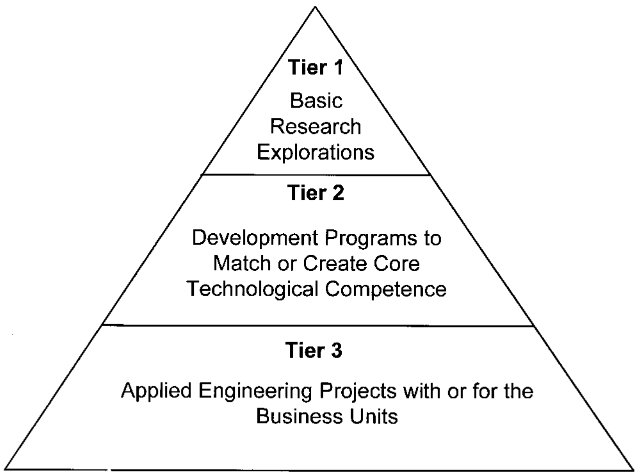
Balancing Market- and Research-Driven R,D&E

Research, development, and engineering (R,D&E) provide the science and technology which firms use to serve tomorrow's customers profitably. Many managers, consultants, and researchers have argued that, to succeed in the next century, R,D&E should be market-driven. See Griffin and Hauser, 1996. John Seely Brown's comments are typical of those heard in interviews with Chief Technical Officers (CTOs) and Chief Executive Officers (CEOs). Indeed, a recent international CTO task force on the evaluation of R,D&E opines that success is more likely if a product delivers unique benefits to the user (EIRMA 1995, p. 36).

However, it is not easy for R,D&E to be market-driven. If we limit our definition of the customer to ''today's customers,'' it might not even be desirable. R,D&E, almost by definition, represents the long-term technological capability of the organization. While many successful new products are developed based on customer needs (von Hippel 1988), an organization cannot meet customer needs if it does not have the capability to do so (EIRMA 1995). The laser was not invented to provide high quality music or to store large quantities of data on compact disks. The U.S. Army Research Laboratory (ARL) and its affiliated research, development, and engineering centers (RDECs) would not have been able to adapt rapidly to the post-cold-war era if they did not have capabilities in the basic research areas. By maintaining basic chemistry and chemical engineering expertise, the Hoechst Celanese Advanced Technology Group, a major producer of chemicals for automotive tires, was able to turn a chance discovery of a chemical process into a thriving pharmaceutical business. Other examples include Carother's research on linear superpolymers that led to nylon and Westinghouse's research on water flows through porous geological formations that led to breakthroughs in uranium mining, the evaluation of environmental impacts for real estate development, and heat flow analyses for high-temperature turbines and for belowground heat pumps (Nelson 1959, Mechlin and Berg 1980). On the other hand, the great isolation of Bayer A.G.'s corporate research center was a failure (Corcoran 1994).

Perhaps today's popular conviction that R,D&E should be evaluated based on market outcomes is too strong. For example, Mansfield (1980) demonstrates that, holding total R,D&E expenditures constant, an organization's innovative output is directly related to the percentage of expenditures allocated to basic research. In a statistical study of new product development at 135 firms, Cooper and Kleinschmidt (1995) find that adequate resources devoted to R,D&E are a key driver that

Figure 1 Tier Metaphor for R,D,&E



separates successful firms from unsuccessful firms. Bean (1995) indicates that a greater percentage of research activities in R,D&E (vs. business units) implies more growth.

We seek to understand how metrics can be used to manage R,D&E more effectively. Specifically, we examine how the use of market-outcome metrics should vary as research activities move from basic explorations to applied projects. We demonstrate how risk, time lags, scope, spillovers, and the management of creative people affect the metrics used to evaluate R,D&E.

Our methodology combines qualitative and quantitative methods. We began by interviewing 43 CTOs, CEOs, and researchers at 10 research-intensive organizations. See Table 1. We next reviewed the public statements of CTOs, consultants, and academic researchers. (See Zettelmeyer and Hauser 1995 for more details on the qualitative interviews and Hauser 1996 for an annotated bibliography.) Together these activities led to both a qualitative description of R,D&E's activities and formal analyses that attempt to generalize the insights. These insights suggest the properties of metrics that can be used to evaluate and manage R,D&E more effectively.

The remainder of this paper is structured into five sections. In the next section we describe a tier metaphor for R,D&E. We then devote a section to each tier. We close with a summary and suggested extensions.

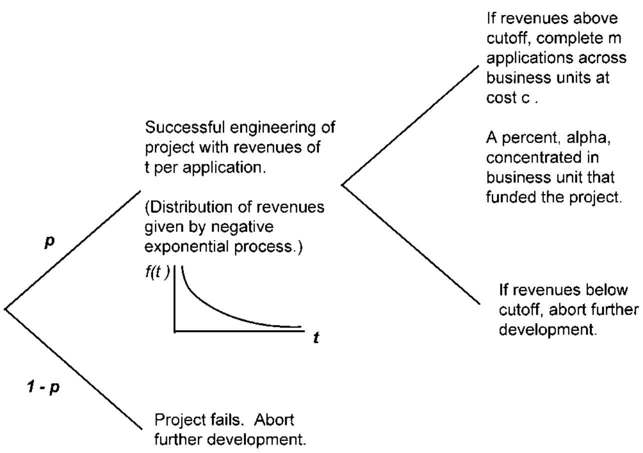
A Tier Metaphor for Describing R,D&E

Many of the firms we interviewed used a tier metaphor to describe their R,D&E activities (Figure 1). This metaphor recognizes that R,D&E activities vary based on risk, on the time lag from conception to market outcomes, and on the number of potential applications (scope). Tier 1 represents basic research. Activities in this area are exploratory and less tied to the market; they concentrate on understanding basic phenomena that might have applicability to many business units. They are often long-term and risky. Tier 2 represents the development of core-technological competence. Tier 2 activities fulfill an organization's existing strategic directions and set new ones. Tier 3 is applied engineering. Activities in tier 3 are usually done with some funding from business units and are often evaluated based on market outcomes. Not only is the tier metaphor common at the firms we interviewed (for example, the U.S. Army uses funding numbers such as 6.1, 6.2, and 6.3 to describe its tiers), but it is consistent with concepts in the R,D&E literature (Bachman 1972, Krause and Liu 1993, Pappas and Remer 1985, Tipping et al. 1995).

Some firms use a formal tier structure, while others use the metaphor to aid evaluation and management. Although many firms assign activities to tiers, all recognize that the assignment is fuzzy. Some activities overlap tiers and most activities evolve from one tier to another as knowledge is gained. Real explorations, programs, and projects often have elements of more than one tier. Indeed, many scientists and engineers work on activities drawn from two or more tiers. We use the tier metaphor to focus on activities that have properties typical of each tier. This metaphor simplifies exposition and makes the insights more transparent. For example, we treat the value of research scope in tier 3 as if it were fully determined by tier 2 activities. In reality, there is still residual uncertainty about research scope that is resolved by tier 3 activities. Thus, the lessons of tier 2 apply to tier 3, but to a lesser extent. By focusing our analyses by tier, we avoid repetition.

We present the tiers in a pyramid to represent conceptually the amount of funding that is allocated to

Figure 2 Decision Tree Representing Project Options



the tiers. For example, in a study of 108 corporations, Mansfield (1981) found that roughly 5% of company-financed research was devoted to tier 1. However, this does not mean that tier 1 is unimportant. In many ways, tier 1 is the research and development (R&D) lab of the R,D&E lab.

In the R,D&E literature, many words—such as *program* and *project*—are used interchangeably (Steele 1988). For the purpose of this paper we adopt Steele's terminology and use the words *objectives* and/or *explorations* for basic research activities, the word *programs* for development activities, and the word *projects* for applied engineering activities. This trichotomy is somewhat arbitrary, but it indicates clearly to which tier we refer.

Tier 3. Applied Engineering for R,D&E's Customers

We begin our analyses with the most market-oriented of the tiers, applied engineering (tier 3). Activities in this tier have the following properties: (1) the business unit managers have the knowledge and skill to evaluate the projects; (2) the projects have more immediate application with relatively less risk; and (3) previous R,D&E activities have provided acceptable estimates of scope, the time stream of payoffs, the magnitude of payoffs, and the probability of success. We focus on metrics that are used to select among tier 3 projects.

Qualitative Ideas

Our interviewees suggested that project selection is the most important and difficult management task in tier 3. They were more satisfied with the monitoring and feedback mechanisms that they used once a project was selected. Many CTOs believed that the business units (the customers of tier 3) have the means and information with which to judge tier 3 projects. Furthermore, they believed that the business units were better able to judge a project's value than R,D&E management. We found a major trend toward making project selection more business-unit driven.

Among the statements that we heard were, "Customer satisfaction is the number one priority;" "R,D&E has to be developed in the marketplace," "The key is to become customer focused;" and "Technology assessment is 'What does it do for the customer?' " At one firm, R,D&E proposes tier 3 projects and the business unit managers decide whether or not to fund them. In many firms R,D&E maintains its budget by "selling" projects to business units.

On the other hand, many firms subsidized R,D&E with central funds. Business units were asked to pay only a fraction of the cost of applied engineering projects. One interviewee stated that the business units could judge research better if they did not have to pay the entire cost. For other examples of subsidies see Corcoran (1994), Mechlin and Berg (1980), and Szakonyi (1990).

Our interviewees proposed at least three justifications for subsidies: research scope, risk aversion, and varying time horizons between the business unit managers and the corporation. Research scope affects subsidies when the results of a pilot test have applications beyond those for which the business unit paid. Other business units often benefit without incurring R,D&E costs. See Mansfield (1982), Mechlin and Berg (1980), and Vest (1995). Scope economies also apply across technological disciplines—for example, when discoveries in chemistry enhance research in biology (Henderson and Cockburn 1996, Koenig 1983). Risk aversion affects subsidies when, without a subsidy, a business unit manager would decide to avoid a risky project even though the expected payoff to the firm justifies the project. Different time horizons affect subsidies when, as expressed in our interviews, business unit managers have shorter time horizons than the firm. They often favor quick fixes for their immediate problems. See Braunstein and Salsamendi (1994), Hultink and Robben (1995), Negroponte (1996), and Whelen (1976). Holmstrom (1989) adds theoretical justification that market expectations can make it rational for managers to be short-term oriented.

Finally, in calculating the net value of an applied project, many firms recognize that they need only commercialize those technologies that prove profitable in pilot tests (Mitchell and Hamilton 1988). The cost of commercialization can be avoided for failed pilot projects. We assume that the firm implements strategies that minimize the tendency of business unit managers to escalate commitments to failing projects (Boulding et al. 1997).

We now incorporate these ideas into a formal model.

Model

We illustrate the contingent nature of applied research decisions with the simple model in Figure 2. First, business unit managers and/or R,D&E managers (and engineers) select among potential projects and begin initial development. For project j, let the pilot engineering costs be k_j. If the project succeeds (with probability p_j), the business unit and R,D&E managers observe the commercial value ($t_j \geq 0$) of the project. This commercial value is modeled as being drawn from a probability density function, $f(t_j)$. If the project fails or if the realized commercial value is below a cutoff (t_c), then the firm can abort the project without further costs. If the commercial value is sufficient, the firm can exercise its ''option'' and apply the technology elsewhere in the firm. We model this *research scope* as if the firm can apply the technology to m_j applications at a cost of c_j for each application. Let α_j be the percent of the applications that are within the business unit that funded the project. (For tier 3 we assume α_j and m_j are given. In the next section, we address how tier 2 might determine these values.)

The parameters in Figure 2 are feasible to obtain. Many organizations routinely make judgments about the expected value of a pilot test ($E[t_j]$), the probability of success for various outcomes (p_j), and costs (both for the pilot application, k_j, and for eventual commercialization, c_j). For example, EIRMA (1995) suggests that the ''3 main components that must be estimated for any project are project cost, benefits, and probability of success.'' See Abt et al. (1979), Block and Ornati (1987), Boschi et al. (1979), Krogh et al. (1988), and Schainblatt (1982) for discussion and methods.

To model the difference in time horizons we define γ_j and γ_F as the business unit and firm discount factors, respectively. These factors reflect the fact that commercial values and costs are really time streams of revenue and costs. If the business unit managers and the firm discount these time streams differently, then the net present value as perceived by the business unit managers will differ from that perceived by the firm. Without loss of generality, we normalize $\gamma_F = 1$ and treat γ_j as the value *relative* to the firm. The business unit manager is more short-term oriented when $\gamma_j < 1$. For issues in the measurement of γ_j, see Hodder and Riggs (1985) and Patterson (1983).

For simplicity, we include all project costs in k_j such that t_j is positive. This allows us to illustrate the effect of $f(t_j)$ with a negative exponential distribution with expected value λ_j. Such probabilistic processes are common in the R,D&E literature. When the business unit managers are risk averse we model them as constantly risk averse with utility, $u(x) = 1 - \exp(-rx)$, where x is monetary outcomes and r is the risk aversion parameter.[1] For risk neutrality, $u(x)$ becomes linear as $r \to 0$.

Analyses

In the appendix, we show that the optimal cutoff, t_c, equals the cost of commercialization, c_j, and that the expected rewards (to the business unit) of the decision tree in Figure 2 are:

$$\text{Expected net rewards} = \gamma_j \alpha_j m_j p_j \lambda_j e^{-c_j/\lambda_j} - k_j. \quad (1)$$

The computations are straightforward applications of conditional probability. The term, $\exp(-c_j/\lambda_j)$, appears in the formula to represent the fact that the firm need only invest further (and incur costs of c_j) when t_j is above the cutoff. The expected outcome from the decision tree in Figure 2 exceeds the naïve valuation, $\gamma_j \alpha_j m_j p_j(\lambda_j - c_j) - k_j$, that would be made without anticipating the option nature of the project.

[1] The qualitative implications should be the same for most reasonable density and utility functions. Some readers may prefer a two-parameter lognormal distribution to facilitate the option-value calculations and to separate risk from expected outcomes.

If the business unit manager is risk neutral, he or she will value the project via Equation (1). If the manager is risk averse, the certainty equivalent (c.e.) can be approximated by:

$$\text{c.e. of expected net rewards} \approx R_j \gamma_j \alpha_j m_j p_j \lambda_j e^{-c_j/\lambda_j} - k_j \quad (2)$$

$$\text{where} \quad R_j = \frac{1}{1 + r\lambda_j m_j \alpha_j}.$$

For risk neutrality, $R_j \equiv 1$. The firm values the project differently than the business unit managers. It earns value from all commercializations within the firm, discounts future value and cost streams less, and can diversify risk. The firm will want at least one business unit to select the project if:

$$m_j p_j \lambda_j e^{-c_j/\lambda_j} - k_j \geq 0. \quad (3)$$

Subsidies

Comparing Equations (1) and (3), we see that the firm can match its incentives with those of the business unit managers by subsidizing projects. If business units are asked to pay only a fraction, s_j, of the project costs, then the business unit manager(s) will choose the same projects as the firm if:

$$s_j = \alpha_j \gamma_j R_j. \quad (4)$$

In other words, the subsidy adjusts for the concentration of research scope (α_j), short-termism (γ_j), and risk aversion (R_j). The subsidy varies by project because both scope and short-termism vary by project. (Short-termism varies because the effect of a differential discount rate has a greater impact on projects with longer time horizons. R_j varies by project because the uncertainty in payoffs varies by project.)

In principle, the subsidy also varies by business unit. Thus, the firm needs a means by which it can entice either a single business unit or a coalition of business units to fund a project. (The firm benefits if other business units "free ride" on the initial business unit's investment. We leave strategic free riding among business units to future papers.)

In theory, the firm can implement the subsidy with a Dutch auction, lowering s_j until one and only one business unit selects the project (with the limit that the subsidy is not so low that Equation (3) is violated). In practice, the subsidies, which vary from 30% to 90% among our interviewees, are set by a complex negotiation process that allows information to be transferred and coalitions to form. (One manager called this *tin cupping* because, like a beggar with a tin cup, she went to other business unit managers asking them to contribute to projects that she championed.) An average subsidy will introduce selection inefficiencies whenever there is substantial variation in α_j, γ_j, λ_j, and m_j.

We summarize this section by stating the implications of Equations (1)–(4) as a set of qualitative hypotheses that can be used for empirical testing. Equations (1)–(3) provide explicit quantification of the value of applied engineering projects.

IMPLICATION 1. (a) *The option value of a tier* 3 *project should be higher than the* (*naïve*) *expected value. This option value anticipates future decisions on subsequent investment.* (b) *For applied projects, firms should use subsidies and implicit auctions. The subsidies and auctions correct for the tendency of business unit managers to choose projects that are more concentrated in a single business unit, have shorter-term payoffs, and are less risky than the firm would find optimal.* (c) *Subsidies should be larger* (s_j *smaller*) *when projects have benefits that are less concentrated, have revenue streams over longer periods, and are perceived as more risky.*

Tier 2. Development Programs to Match or Create Core Technological Competence

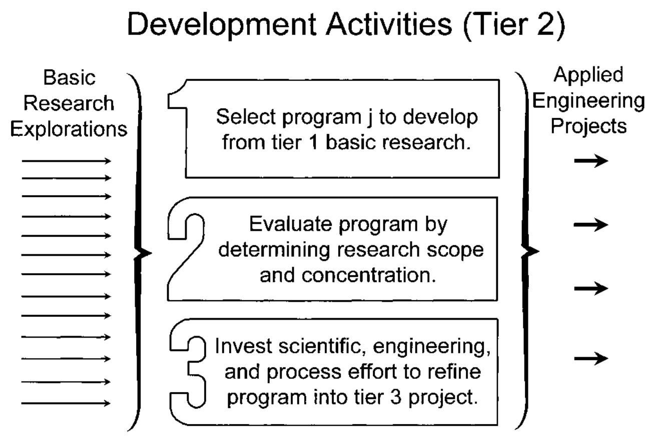
We now focus on development activities (tier 2) that provide the bridge from basic research (tier 1) to applied engineering (tier 3). These activities are more risky and have longer-term payoffs than tier 3 projects. They are also more difficult for business unit managers (and line managers) to evaluate because evaluation requires more detailed information and greater current technical experience. Instead, business unit managers rely more heavily on the decisions of R,D&E managers and engineers. The challenge for activities having tier 2 characteristics is to develop a set of metrics with which to evaluate the decisions and the efforts of R,D&E managers and engineers. Because the firm must rely on their decisions, we seek metrics that encourage R,D&E managers and engineers to make those decisions and allocate those efforts that are in the firm's best interests. In addition, because tier 2 programs evolve into tier 3 projects, we examine how the activities in tier 2 determine the parameters used to select tier 3 projects.

Qualitative Ideas

Our qualitative interviews and the R,D&E literature suggest that the primary task of development is to match expertise with strategic direction. See Adler et al. (1992), Allio and Sheehan (1984), Block and Ornati (1987), Boblin et al. (1994), Chester (1994), EIRMA (1995), Frohman (1980), Ransley and Rogers (1994), Schmitt (1987), Sen and Rubenstein (1989), and Steele (1987, 1988). As one of our interviewees said, "The customer knows the direction, but lacks the expertise; researchers have the expertise, but lack the direction." Tier 2 researchers and managers are judged both for their competence in developing technologies and for their ability to align the values of R,D&E with those of the firm (Steele 1987). Our interviewees said that development succeeds if it gets the programs right. However, researchers in tier 2 must also have the incentives to invest the right amount of scientific, engineering, and process effort.

R,D&E researchers (and managers) appear to have more expertise and knowledge than top-level managers about the specifics of the development programs. Thus, firms use metrics to encourage tier 2 researchers to select the right programs and to put forth sufficient scientific, engineering, and process effort to develop those programs. We heard concerns that net present value metrics favor short-term, predictable, incremental development programs (Steele 1988, Irvine 1988). Our interviewees believed that tier 2 metrics should not imply a penalty for failure that is too strong. Such penalties encourage researchers to focus only on safe technologies and not take sufficient risks. Failure was part of the territory (estimates of failure ranged from 20% to 80%); interviewees felt that metrics that eliminated failure also eliminated success. Instead, we often found metrics such as patents, publications, citations, citations to patents, and peer review. See also Edwards and McCarrey (1973), Henderson and Cockburn (1996), Irvine (1988), Miller (1992), Pappas and Remer (1985), and Shapira and Globerson (1983). These metrics appear to be surrogates for the scientific, engineering, and process *effort* that is devoted to development programs. There appears to be a tension, when designing a tier 2 evaluation system, between market-outcome metrics and *effort-indicator* metrics.

Figure 3 **Representation of Development Activities**



Model

Figure 3 represents our conceptual model of tier 2 activities. In step 1, researchers select programs based on the ongoing results of basic research (tier 1) explorations.[2] Naturally, tier 2 researchers do so anticipating potential outcomes but taking uncertainty into account. In step 2, researchers evaluate each program to resolve some of the uncertainty. In this evaluation they determine research scope (m_j) and concentrations (α_j's for each business unit). This step also clarifies uncertainty in the value (to the firm) of the program so that business unit managers and applied engineers have sufficient information to estimate the parameters for Equations (1)–(4). If the program shows sufficient potential, then, in step 3, development researchers invest significant scientific, engineering, and process efforts to develop the program into potential applied projects.

Because development researchers select programs before they know the outcomes of the development programs, we model a key parameter, *research scope*, as a random variable, $\tilde{m}_j$. Specifically, we model the process of determining $\tilde{m}_j$ as if there were M_j potential applications within the firm. During step 2, the researcher

[2] We refer to development decisions and efforts as if they were made by researchers. The same analyses apply to teams of researchers and managers (as long as we account for free riding within teams).

determines how many of these applications apply to the firm—a priori each applies with a probability, q_j. (Estimates of M_j and q_j are based on the result of basic research explorations and on expertise in evaluating the outcomes of these explorations.) We define v_j as the *value* of each realized application.

We model the scientific, engineering, and process effort in step 3 with an additive parameter, e_j, that measures the expected incremental profit to the firm of this effort.[3] The effort by development researchers to obtain these results is costly to the researchers and this cost may be difficult for the firm to observe. We call this cost, $d_j(e_j)$, and assume that it is convex in e_j. Finally, there is some fixed cost to the firm, K_j, of developing program j.

Each program might have different anticipated time streams of net revenues and development researchers might be more short-term oriented than the firm. We model this by a discount factor, Γ_j. We allow researchers to be (constantly) risk averse. (We expect that $\Gamma_j < \gamma_j$ because of longer time lags associated with development programs. The case of no short-termism is represent by $\Gamma_j = 1$ and the case of risk neutrality is represented by $r \rightarrow 0$.)

To focus on key phenomena and avoid redundancy, we have simplified our model in this section. Each of these simplifications can be relaxed readily. First, we set $k_j = 0$ to simplify the options analysis that has already been discussed. (Options analysis applies to tier 2 in the same manner that it applies to tier 3.) Second, we model uncertainty in $\tilde{m}_j$ but not v_j because the effect of uncertainty in v_j would only reinforce the effects due to $\tilde{m}_j$. Finally, we model the effort allocated in step 3 but not the effort allocated in step 2. The basic intuition would be the same, but the algebra would be unnecessarily complicated. None of these simplifications change the basic insights derived here.

Development Metrics

Recently, many firms have adopted development metrics that are based on comparing market outcomes to development costs. For example, see McGrath and Romeri (1994). However, some of our interviewees believe that such schemes distort development decisions. Thus, we want to contrast these metrics with effort-indicator metrics.

Because many firms try to measure effort directly with metrics such as publications, citations, patents, citations to patents, and peer review, we represent these metrics with a normal random variable, $\tilde{e}_j$, with mean e_j and variance, σ_e^2. The uncertainty in this measure represents the fact that these metrics are, at best, noisy indicators of the incremental profit to the firm of the researchers' efforts.

Market outcomes result from the value and scope of the chosen program and from the researchers' efforts. To explore development metrics, we recognize that the market outcomes in our model are $\tilde{m}_j v_j + \tilde{e}_j$ and the costs are K_j. This implies a net market-outcome metric of $\tilde{m}_j v_j + \tilde{e}_j - K_j$. To represent our observations that firms combine market-outcome and effort-indicator metrics, we consider a more general metric that allows a weight of η_1 on market outcomes, η_2 on effort, and η_3 on costs. If we define $\beta_v = \eta_1$, $\beta_e = \eta_1 + \eta_2$, and $\beta_K = \eta_3$, then this implies the linear development metric given by Equation (5).

$$\text{Development metric} = \beta_v \tilde{m}_j v_j + \beta_e \tilde{e}_j - \beta_K K_j. \quad (5)$$

In this notation, the metric advocated by McGrath and Romeri is represented by a special case where $\beta_v = \beta_e = \beta_K = 1$, or equivalently, $\eta_1 = \eta_3$ and $\eta_2 = 0$.[4] The linear function suffices to demonstrate the basic tension in development metrics. However, future analyses might improve on observed practice by introducing nonlinear reward systems.

Development metrics enable top management to motivate researchers to choose those development

[3] We define e_j based on effort that is induced incrementally by the metrics system above and beyond any effort that the researcher would put forth based solely on his or her base wage. We might consider alternative formulations treating either e_j or $(1 + e_j)$ as multiplicative terms. These formulations provide the same qualitative implications when we focus on program choice or effort allocation. However, scaling constants and the detailed optimizations vary.

[4] Specifically, their effectiveness index (EI) is equal to (% of revenue from new products)*[(% of revenue that is profit)/(% of revenue spent on R&D) + 1]. For clarity of exposition our representation is a linear rather than a ratio function. We might also note that their metric does not include the impact of development activities on existing products.

programs (and allocate effort) that are in the best interests of the firm. Top management would have less need for these metrics if it could simply dictate to the researchers the programs on which they should work and then monitor costlessly how hard they work. The metrics enable top management to delegate the selection of programs and the allocation of scientific, engineering, and process effort to those who have the unique technical knowledge and experience necessary to judge the merits of the programs.

To represent how researchers will evaluate rewards, explicit or implicit, that are based on this metric, we first recognize that researchers will find effort to be costly. Thus, we subtract $d_j(e_j)$. Secondly, we recognize that there is a time lag in observed outcomes, but not costs. Thus, we discount observed outcomes. Furthermore, if $\tilde{e}_j$ is observed before $\tilde{m}_j v_j$, then we allow different discounting constants, Γ_j^m and Γ_j^e.[5] Finally, if researchers are risk averse they will perceive the uncertainty in $\tilde{m}_j$ and $\tilde{e}_j$ to be costly. Thus, we represent the uncertain rewards with their certainty equivalent. In the appendix we derive the researcher's certainty equivalent based on the development metric:

$$\text{c.e.} = \beta_v \Gamma_j^m M_j q_j v_j + \beta_e \Gamma_j^e e_j - \beta_K K_j - d_j(e_j) - (r/2)\{\beta_v^2 (\Gamma_j^m)^2 M_j q_j (1 - q_j) v_j^2 + \beta_e^2 (\Gamma_j^e)^2 \sigma_e^2\}. \quad (6)$$

It is immediately clear that either β_v or β_e must be nonzero. Otherwise, researchers would select no programs for development and allocate no effort.

In contrast to researchers, the (risk neutral) firm wants to select those programs that maximize the *expected value* of the program (net of the wages the firm must pay). To calculate this value we use standard agency theory methods (e.g., Holmstrom 1989) to represent the profit the firm can earn. First, we recognize that $M_j q_j$ is the expected value of $\tilde{m}_j$ and e_j is the expected value of $\tilde{e}_j$. Thus, before wages, the firm's expected profits are $M_j q_j v_j + e_j - K$. However, if the firm is to retain its employees, it must pay them their market wages net of switching costs, w_o, *and* it must reimburse them for any effort costs and for any risk costs. (By definition, w_o represents the minimum amount that would be required to retain a researcher who did not have to incur incremental effort and risk costs on the firm's programs.) Thus, the firm's profit is given by:

$$\text{Firm's profit} = M_{j^*} q_{j^*} v_{j^*} + e_{j^*}^* - K_{j^*} - d(e_{j^*}^*) - w_o - \text{risk costs}, \quad (7)$$

where j^* indicates the researchers' program selection and $e_{j^*}^*$ indicates the researchers' response to the firm's choice of the β's. The firm will select the β's to maximize its profit. This optimization will, by implication, determine the program choice and the effort that the researchers allocate.

In principle, we could solve the complete agency problem by choosing the β's to maximize Equation (7), recognizing that the certainty equivalent of the researchers' wages is given by Equation (6). The resulting solution would balance the tension between inducing the best choice and motivating the optimal effort. However, we gain greater insight into this tension with a simpler approach that analyzes the problem in stages. We begin by holding effort constant and illustrating how β_v and β_K affect the choice among programs. We then hold research scope constant to show how β_e affects the researchers' efforts. This allows us to interpret the relative magnitudes of β_v, β_K, and β_e.

Selecting the Right Programs

For this subsection, we assume that e_j^* and $d(e_j^*)$ do not vary by research program and that $\sigma_e^2 = 0$. Under these conditions, the anticipated effort allocation will not affect program choice.[6] With efforts constant among programs, the effort benefits, effort costs, and fixed wages would simply shift K_j by a fixed constant in the following discussion. Thus, we can normalize $[e_j^* - d(e_j^*) - w_o] = 0$ without loss of generality.

Differential discounting ($\Gamma_j^m < 1$) and risk aversion ($r > 0$) cause the researcher's c.e. to differ from the expected profit the firm could earn if it did not need to rely on metrics and could dictate the choice of program.

[5] We have chosen to define the Γs with respect to the βs rather than the ηs. It is possible to derive one from the other.

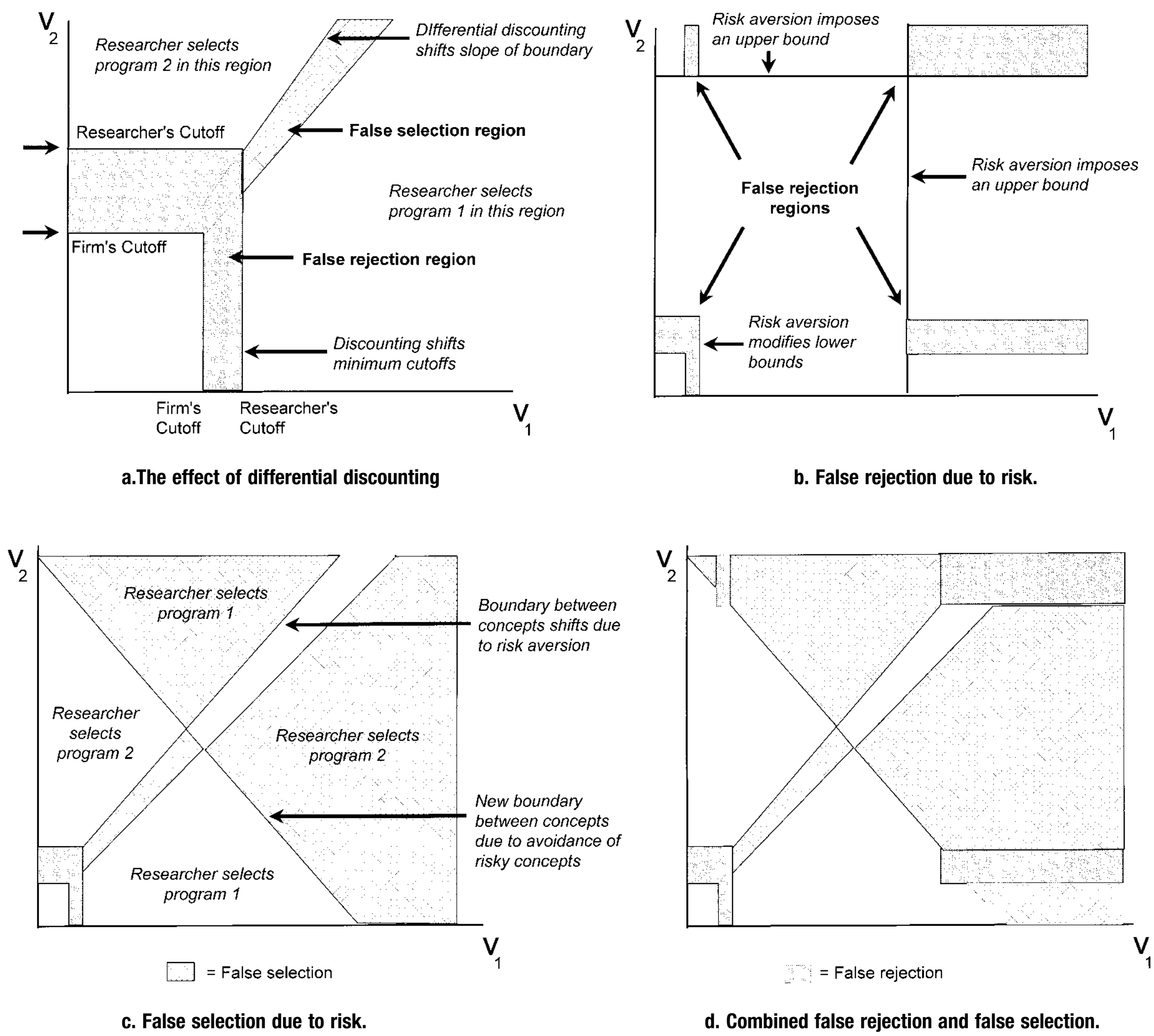
[6] The technical conditions of the problem formulation assure us that we can choose β_e^* independently of β_v^* and β_K^*. Thus, all terms involving e will be the same for each project being compared. For a multiplicative formulation, e_j would scale the value and e_j^2 would scale the variance. If $e_j = 1$ for the multiplicative formulation, then Figure 4 would be the same.

In the latter case the firm's profit would be $M_j q_j v_j - K_j$. For program choice to matter to the researcher, Equation (6) requires non-zero β_v. However, larger β_v increases the firm's risk costs. Furthermore, Equation (6) suggests that β_v could distort program choice. That is, differential discounting and risk aversion might cause researchers to reject some programs that would be profitable for the firm and to favor less profitable programs (for the firm) over more profitable programs.

We find it is easier to illustrate these effects graphically. Figure 4 maps the magnitude of the phenomena for the case of two alternative research programs and for representative values of the parameters (given in the appendix). The horizontal and vertical axes represent the values (v_j) of programs 1 and 2, respectively.

Figure 4a isolates the effect of discounting (with risk neutrality). Equations are derived in the appendix. If researchers discount the time stream of revenue, then some programs will be falsely rejected (inverse L-shaped region in Figure 4a). If revenues from one research program occur faster than another ($\Gamma_1^m > \Gamma_2^m$), then researchers will be more likely to choose the program with better short-term prospects (diagonal false selection region in Figure 4a). We can eliminate the false

Figure 4 How Development Metrics Affect Program Choice



rejection regions if $\beta_K = \Gamma_j^m \beta_v$, but eliminating the false selection region requires, in addition, that we allow β_v to vary by program.

Figure 4b isolates the effect of risk on false rejection. (We expand the scale in Figure 4b, vs. Figure 4a, to illustrate this effect.) When $\tilde{m}_j$ is a random variable and researchers are risk averse, the certainty equivalent will be less than the expected value (see also Holmstrom 1989). For a given cost (K_j), when the value (v_j) and implied risk become large, the certainty equivalent becomes negative and researchers no longer find it attractive to begin development even though the program provides a very large expected return to the firm. The areas where *both* programs are falsely rejected are shaded. (We might also shade the regions above the upper bound to illustrate that at least one program is falsely rejected.) If the firm wants to eliminate these false rejection regions, it must make β_v sufficiently small, such that the false rejection regions are beyond any feasible outcome, but large enough so that researchers prefer high-expected-return programs. Placing too large a weight on market-outcome metrics leads to a tendency by researchers to avoid high-expected-return development programs that are risky and/or long-term.

Figure 4c isolates the effect of risk on false selection. The concept is similar to that of false rejection. In the shaded regions of Figure 4c, uncertainty and risk aversion cause researchers to avoid high-return development programs when the returns are risky and/or long-term. The firm can eliminate these false selection regions by making β_v sufficiently small.

Figure 4d summarizes the effects of both discounting and risk. The regions are more complex, but the phenomena are the same—discounting and risk aversion lead to large regions of false rejection and false selection when researchers are evaluated too heavily on market-outcome metrics.

Encouraging Tier 2 Scientists and Engineers to Put Enough Effort into Developing a Program

In this subsection we focus on the effort that is allocated after a program is selected. We hold the realized scope ($\tilde{m}_j$), the value (v_j), and costs (K_j) constant and focus on step 3 in Figure 4. With only effort being analyzed, the selection of a weight (β_e) to encourage researchers to allocate optimal efforts is a standard agency theory problem. See Holmstrom (1989). In the appendix we show that the firm can choose an *optimal* β_e such that researchers allocate the scientific, engineering, and process effort that maximizes the firm's profits. The optimal weight is:

$$\beta_e^* = (\Gamma_j^e)^{-1}\left[1 + r\sigma_e^2 \frac{\partial^2 d(e_j^*)}{\partial e_j^2}\right]^{-1}. \tag{8}$$

Because $d(e_j^*)$ is convex, $\Gamma_j^e \beta_e^* \in [0, 1]$. When researchers are very good at anticipating the outcomes of their efforts, σ_e^2 will be close to 0.0. When the effort metrics are observed much faster than market outcomes, Γ_j^e will be close to 1.0. Under these conditions, β_e^* will be close to 1.0.

We now see the tension. If market outcomes were the only metrics available, then the metrics would measure $\tilde{m}_j v_j$ and $\tilde{e}_j$ simultaneously. To avoid false program *choice* the firm would want the weight on market outcomes to be small, but to induce the right research and process *efforts* the firm would want the weight on market outcomes to be large. One way to finesse this tension is for the firm to search for metrics that correlate with effort, but not necessarily with market outcomes. The firm can then implement a small weight on $\tilde{m}_j v_j$ and a large weight on $\tilde{e}_j$ by placing a small weight on market outcomes and a large weight on the effort-indicator metrics.[7] The firm finds it attractive to use effort-indicator metrics more than market outcomes because (1) effort-indicator metrics can be observed sooner than market outcomes and (2) the measurement uncertainty relating the effort-indicator metrics to true effort is less than the uncertainty in predicting ultimate market outcomes. The reduced discounting and risk motivate researchers to allocate the most profitable amount of effort to the development programs. The effort-indicator metrics make it feasible for the firm to place a small, but positive, weight on market outcomes. A small weight on market outcomes avoids false selection and false rejection in the choice of development programs.

Selecting the Right Programs *and* Allocating Sufficient Effort

If returns to effort vary by development program, then, in step 2 of Figure 4, for a given set of βs, researchers

[7] Returning to the ηs for a moment, we see that β_v small and β_e large imply η_1 small and η_2 large, and vice versa.

will select among programs anticipating the effort that they will allocate in step 3. Technically, we incorporate this effect by using all the details of Equations (6) and (7) to redo the analyses that led to Figure 4 and Equation (8). For each potential program, these optimal values of effort do not depend upon the realized value of the research scope because, in our model, $\tilde{m}_j$ and $\tilde{e}_j$ are independently distributed (and researchers are constantly risk averse). More complete analysis could determine the optimal metrics (βs).[8] However, this more complicated analysis does not change the qualitative lessons that can be derived from our simpler analyses.

Implications for Practice

Our simple analyses seem to conform to practice. Development metrics do appear to be based on both market-outcome metrics and effort-indicator metrics. In particular, many firms use metrics such as patents, publications, citations, citations to patents, and peer review. Such metrics have proven to be correlates of incremental value, and by implication, scientific, engineering, and process effort. See Griliches (1990), Koenig (1983), Miller (1992), Stahl and Steger (1977), and Tenner (1991). Indeed, if more than one such measure of effort is available, the firm can do better by using a linear combination of measures (Holmstrom 1989). When the measures are independent indicators, the optimal weights are inversely proportional to the variance of the measures (see the appendix for equations). Thus, when metrics can be found that are indicators of development effort, the firm should weigh these metrics more heavily than market-outcome metrics. If these indicators can be observed before market outcomes ($\Gamma_j^m < \Gamma_j^e$) and if the measures are less uncertain from the perspective of development researchers, then effort-indicator metrics help to avoid distortions due to short-termism and risk aversion.

Our analysis is contrary to calls in the popular press for greater market accountability of development and is contrary to many of the schemes advocated (but not yet fully evaluated) in the R,D&E literature. We predict that a simple comparison of market outcomes and research costs (e.g., McGrath and Romeri 1994) will lead researchers to avoid long-term and/or risky programs. (Indeed, one senior manager, who indicated to us that his firm uses these measures, found that the measures increased for a few years, but now appear to be decreasing.)

In addition to combining market-outcome and effort-indicator metrics, the firm can also attempt to develop metrics that measure directly the ability of researchers to choose the right programs. For example, some firms reward development researchers for ''strategic vision'' and for decisions that are aligned with the firm's goals (Steele 1987).

We summarize our analyses with some testable implications.

IMPLICATION 2. *Development programs* (*tier* 2) *should be evaluated on market outcome metrics such as profits, revenues, sales, or business-unit evaluations, but the weight on those metrics should be small. Otherwise, researchers favor short-term programs with less risk. On the other hand, metrics such as publications, citations, patents, citations to patents, and peer review should have a much higher weight* (1) *if these metrics correlate with the amount of value-enhancing scientific, engineering, and process effort and* (2) *if they can be observed sooner and with less uncertainty than market outcomes.*

Tier 1—Basic Research Explorations: The Role of Research Tourism

We now focus on basic research explorations (tier 1) that provide the raw material for development programs. The uncertainty and time lag for these explorations is even larger than that for development programs and line managers must rely even more on the specialized knowledge of tier 1 managers and researchers. Many of the lessons from previous tiers apply to tier 1. For example, effort-indicator metrics should be given a higher weight than market-outcome metrics. The additional challenge in tier 1 is to provide the right incentives so that tier 1 researchers and managers explore a sufficiently broad set of new ideas, concepts, technology, and science.

[8] The profits that result from optimal βs will be less that the (''first-best'') profits the firm could obtain if it had the knowledge and capabilities to dictate program choice. The metrics-based profits are less because the firm must reimburse the researcher for the risk costs that the development metrics impose. Future authors might reduce the risk costs with a nonlinear system to obtain ''second-best'' profits. (Optimization over all potential linear or nonlinear functions.)

Qualitative Ideas

We found that basic research (tier 1) is more likely than the other tiers to be funded from corporate coffers; more likely to be located in central laboratories; and more likely to focus on long-term concepts. (One of our interviewees, the CEO of a \$2 billion company, said that one of his main responsibilities was to protect the basic research budget from his business unit managers.) See also Chester (1994), Krause and Liu (1993), Mansfield (1981), Mechlin and Berg (1980), Reynolds (1965), and Szakonyi (1990). Tier 1 is organized more often by scientific discipline than by markets served (see also Chester 1994). It accounts for roughly 5% to 15% of R,D&E spending, but appears to be the seed for new ideas, concepts, technology, and science.

Our interviewees stressed the need to maintain the best, most creative basic researchers (see also Steele 1988). We observed that management provided these people with sufficient protected space and discretion in which to innovate. This included special privileges, such as ''Research Fellows'' at IBM and 3M or ''Man on the Job'' at the U.S. Army, that are not unlike the tenure system at research universities. However, judging the best people was difficult because the success of a research exploration depends, in part, on as-yet-undiscovered natural phenomena. Indeed, some researchers provide value to the firm by identifying which directions *not* to explore. As a result, basic researchers are often judged by the quality of the research that they, themselves, perform (Platt 1964). Fame, recognition, and salary appear to depend more on that which a researcher originates than on ideas, concepts, technology, and science that are ''arbitraged'' from outside sources.

In contrast, many of the most profitable new ideas, concepts, technology, and science come from outside the firm. Our interviewees stressed the need to maintain expertise in the scientific disciplines in order to identify ideas from universities, from other firms in the industry, and from other industries. They called this activity *research tourism*. One of our interviewees stressed that *his* firm's competitive advantage was to identify and develop outside ideas better than anyone else in the industry. Research tourism opens ''new fishing grounds'' for corporate development (Griliches 1990) and spillovers can be quite large (Acs et al. 1992, Bernstein and Nadiri 1989, Griliches 1992, Jaffe 1989, Ward and Dranove 1995). In an econometric study of 1,700 firms, Jaffe (1986) suggests that, while the direct effect of R,D&E spending by competitive firms lowers profitability, the indirect effect of spillovers is sufficiently large to make the net effect positive.

However, encouraging research tourism is not easy. A common problem at many research laboratories is a *Not-Invented-Here* (*NIH*) attitude (Griffin and Hauser 1996). The outputs of internal explorations are easier to measure, hence it is tempting to evaluate researchers based on that which they originate rather than the total number of ideas, concepts, technology, and science that they bring into the firm. This is perpetuated by evaluation systems (e.g., Galloway 1971) that trace successful new products back to their idea source. Other firms encourage work within the organization to avoid ''buying'' technological results (Roussel et al. 1991). EIRMA (1995) suggests that the inability to incorporate spillovers and spin-offs appears to be one of the weaknesses of the evaluation systems used by European firms.

The Right Reward System Encourages Research Tourism; the Wrong Reward System Encourages ''Not Invented Here''

We focus on how the firm should evaluate researchers so that they have incentives to seek out the right amount of ideas, concepts, technology, and science. For ease of exposition, we refer to these outputs simply as *ideas*. (Previously, we addressed how researchers and managers chose which idea to develop as a tier 2 program.) By *right amount*, we seek the number of ideas that maximizes the value of the ideas minus the cost of obtaining them. Some ideas are better than others, but for the purpose of this section we treat all ideas equally.

Our interviews and the literature (e.g., Cohen and Levinthal 1989) suggest that more and better internal research provides a greater ability to identify and use outside ideas. Let h be the number of *internal explorations* and assume that each exploration leads to an *idea*. Suppose that for each internal idea identified, the basic researcher can also identify μ ideas from the outside. Thus, the total number of ideas, n, will be equal to $h + \mu h$. Let κ_i be the cost of exploring an internal idea and let κ_o be the cost of exploring each external idea. (The subscripts are mnemonic for inside and outside, respectively.) Naturally, $\kappa_i > \kappa_o$. Let $V(n)$ be the value of n total ideas (appropriately discounted). We assume that V is a concave function of n. For example, $V(n)$ might be the maximum of n draws from a normal distribution. (The effects of risk and differential discounting on V will be similar to those covered in the previous section. In this section, we focus on the implications of choosing either n or h as the metric. Therefore, we treat $V(n)$ as if it imposes no risk and no time lag on the researchers.)

The potential for spillovers ($\mu > 0$) decreases the cost per idea, hence, for concave V, the optimal number of ideas increases when spillovers are possible. However, even though spillovers make internal explorations more efficient, this efficiency might imply fewer internal explorations. In the appendix we show formally that this means that the optimal number of internal explorations might actually decrease. We summarize this analysis as testable implications.

IMPLICATION 3. *When spillovers are possible*, (*a*) *the optimal number of explorations increases but* (*b*) *the optimal number of internal explorations might decrease*.

Implication 3 suggests why tier 1 researchers might adopt an NIH attitude. If a researcher's (or research manager's) status is based on the number of the internal explorations that the firm funds, then seeking spillovers might decrease this internal empire. To illustrate the phenomenon more formally, suppose that the firm can evaluate researchers on either internal ideas alone (the size of the research "empire") or on the total number of ideas that are identified—whether or not they originate internally. That is, the firm evaluates tier 1 researchers based either on h or on n. We call these evaluation functions $g_h(h)$ and $g_n(n)$. Suppose that the researcher's rewards, either explicit or implicit, are based on these evaluations.

Tier 1 researchers can choose whether or not to seek spillovers. We model this ability by allowing them to choose how many external ideas they explore. That is, they choose a value μ^o from the set $[0, \bar{\mu}]$ such that the total number of ideas they explore is $h + \mu^o h$. Let μ^* be the value they choose (in their own best interests). If $\mu^* = 0$ then this is equivalent to NIH; if $\mu^* = \bar{\mu}$, then this is equivalent to research tourism.

We now examine how the choice of metric affects the researchers' reactions to the evaluation system. To make the comparison meaningful, we select functions such that the researcher would earn the same reward whenever he or she acts in the best interests of the firm. We choose evaluation functions that accurately reflect the value to the firm of the ideas that the researcher explores. These assumptions imply that $g_h(h) = V[(1 + \bar{\mu})h]$ and $g_n(n) = V(n)$. The firm would choose this $g_h(h)$ if it fully expected researchers to explore spillovers and rewarded them accordingly, but did not anticipate that the choice of a research metric affects the researchers' choice of μ^o. (We might also assume that the firm can anticipate the value of μ^o that researchers will choose. If the firm were restricted to using h, but could anticipate μ^o it would choose $g_h(h) = V(h)$; if it were allowed to use n, it would choose $g_n(n) = V(n)$ as the reward function. We obtain similar results for these assumptions.[9])

The formal results are derived in the appendix. We provide the intuition here. When researchers are evaluated on the metric, n, the evaluation structure for researchers is similar to that by which the firm evaluates its profits. The cost per idea decreases with μ^o, thus researchers, like the firm, will find it in their own best interests to set $\mu_n^* = \bar{\mu}$. Their objectives will parallel those of the firm and they will choose the optimal number of explorations. However, when the researchers are evaluated based on the metric, h, the cost per unit gain in $g_h(h)$ increases as μ^o increases, hence the researchers will want to keep μ^o small. With $\bar{\mu} > 0$ and $\mu_h^* = 0$, researchers are rewarded as if there *were* spillovers, but they incur costs as if there *were no* spillovers. Because rewards are concave, this leads to more internal explorations. However, it does not necessarily imply more ideas. That depends upon the relative costs of internal

[9] We could analyze this as a formal agency problem, in which case, the firm could obtain maximal profits by paying tier 1 researchers via $V(n) + w_o + (\kappa_i + \kappa_o\mu)/(1 + \mu)n^* - V(n^*)$. Because we have abstracted from risk in this section (it is covered in previous sections), this makes tier 1 researchers the residual claimants. Alternatively, we could restrict the firm to rewards of the form $g(h)$ + constant. In this case, the optimal rewards would be $g(h) = V(h)$. This case is analyzed in the appendix. It provides similar, but not identical, results. In the text we have chosen to compare the two reward systems that we feel represent practice. We leave analysis with risk aversion and differential discounting to future extensions.

and external explorations. We state these testable results as Implication 4.[10]

IMPLICATION 4. (*a*) *If tier* 1 *researchers are evaluated on all ideas, new concepts, new technology, and new science, including that identified outside the firm, they will set* $\mu_n^* = \bar{\mu}$ *and invest in the optimal number of explorations for the firm*. (*b*) *If researchers are evaluated on the results of internal explorations only, they will adopt an NIH attitude by setting* $\mu_h^* = 0$. *They will work on more internal explorations and may develop fewer ideas, new concepts, new technology, and new science than would be optimal for the firm*.

Summary and Implications for Basic Research Metrics

Our analysis of spillovers suggests that the common practice of rewarding basic researchers for original ideas leads them to (1) ignore ideas that were "not invented here" and (2) build "research empires" by undertaking too many internal explorations. This may lead to fewer ideas. The firm can be more profitable if it encourages research tourism by evaluating researchers for ideas generated internally *and* for ideas identified from sources outside the firm. Fortunately, progress is being made. The recent vision statement adopted by General Motors includes the phrase "Develop more highly valued innovations, *no matter their source*, than any other enterprise." (Emphasis added. Vision statement obtained by private communication to the author.)

Summary and Future Research

Arthur Chester (1995), Senior Vice President for Research and Technology for GM Hughes Research Laboratories, states that "measuring and enhancing R&D productivity or R&D effectiveness . . . has gained the status of survival tactics for the R&D community." R,D&E evaluation is an important policy issue in Japan (Irvine 1988) and Europe (EIRMA 1995). Erickson and Jacobson (1992) provide evidence that there are no supranormal returns to R&D spending, but that "obtaining a comparative advantage . . . depends crucially on the specific nature of the expenditure and how it interacts with the firm's asset and skill base." CEOs and CTOs use metrics to evaluate and manage people, objectives, programs, and projects. In many ways, metrics determine whether or not a firm's R,D&E activities are well managed. While the identification of specific measures for each firm is an empirical question beyond the scope of this paper, we have attempted to identify the properties of those metrics that enable firms to manage R,D&E effectively.

First, it is clear that metrics must vary by tier. Market-outcome metrics make sense for applied engineering projects that provide relatively predictable and immediate returns. However, the incentives of business-unit managers may not be aligned with those of the firm. Thus, the cost of applied projects should be subsidized to adjust for short-termism, risk aversion, and scope. Ideally, these subsidies should vary by project and by business unit.

For development programs with longer-term and more-risky payoffs, market-outcome metrics should be given less weight. This is contrary to popular wisdom. Indeed, too great a stress on market-outcome metrics will encourage managers and researchers to avoid long-term, risky programs that have high profit potential. Instead, for development programs, the firm should place a small weight on market outcomes and a larger weight on effort-indicator metrics such as publications, citations, patents, citations to patents, and peer review. This combination of metrics provides managers and researchers with the incentives to choose the right programs and allocate the right amount of value-enhancing scientific, engineering, and process effort.

Basic research is even further from the market, hence more difficult for line managers and business unit managers to evaluate. As a result, firms rely more heavily on the judgment of basic research managers and scientists. They often seek indicators of the quality of these people and the quality of their work. Unfortunately, many organizations evaluate these people based only on the ideas, concepts, technology, or science that they originate. Such evaluations encourage them to do only *internal* explorations and build research empires that are too large. A firm can do better by encouraging *research tourism*. It should reward research managers and scientists for the ideas that they originate *and* for the ideas that they identify from outside the firm. Problems with "not invented here" result from the wrong evaluation

[10] If tier 1 researchers are evaluated on $g(h) = V(h)$, then the equivalent result is that researchers will develop fewer ideas and may work on fewer internal explorations.

system. They can be avoided with the right evaluation system.

Our analyses can be extended. For example, Draper Laboratories (Pien 1997) has begun to use these insights to identify a set of metrics that provides researchers and managers with the right incentives throughout the tiers of R,D&E. Pien's metrics also show promise in predicting the success of tier 1 explorations. We are extending these analyses to other corporations. We have begun other research to test whether metrics affect people in the ways predicted in this paper. For example, we are sending questionnaires to CTOs to determine whether the metrics used by a cross-section of organizations have the predicted properties.

Other research directions include the integration of R,D&E metrics with internal customer evaluation systems and/or customer satisfaction measures (Hauser et al. 1994, 1996), the exploration of self-selection on risk aversion (Holmstrom 1989), strategic behavior to withhold information or support (Rotemberg and Saloner 1995), internal patent systems and research tournaments (Taylor 1995), product platforms (Utterback 1994), and the role of R,D&E as a crucible for growing technical managers.

Finally, there are personal and cultural issues in a research community. Many scientists are driven by an inherent need to know and many scientists believe strongly in a research culture. Hopefully, our analyses are complementary to these sociological and anthropological approaches to R,D&E management.[11]

[11] This research was funded by the International Center for Research on the Management of Technology (ICRMOT) and the Center for Innovation in Product Development (CIPD), M.I.T. We wish to thank the managers, scientists, and engineers who donated their time to talk to us about this important topic. This paper has benefited from presentations at the M.I.T. Marketing Group Seminar, Stanford University Marketing Seminar, the U.S. Army Soldier Systems Command, the ICRMOT sponsors meeting, the Marketing Science Conference at the University of Florida, the Marketing Science Institute's conference on Interfunctional Issues in Marketing, and at the University of California at Los Angeles. Special thanks to Florian Zettelmeyer who completed most of the qualitative interviews described in this paper. I have discussed many of these ideas with him and have benefited greatly from his feedback. Complete listings of related working papers are available at http://web.mit.edu/icrmot/www/ and at http://web.mit.edu/cipd.

Appendix: Derivations and Proofs

For ease of exposition, we drop the j subscript in the derivations when there is no ambiguity. We assume that all functions are thrice differentiable and, when appropriate, all maxima are interior.

Tier 3: Applied Engineering for R,D&E's Customers

Equation (1). Following the decision tree in Figure 2, we obtain

$$\text{expected net rewards} = (\gamma[(1-p)\cdot 0 + p\cdot \text{Prob}\{t < t_c\}\cdot 0 + p\cdot\text{Prob}\{t \geq t_c\}\cdot\{\alpha m(E[t \mid t \geq t_c] - c)\}]$$

minus the costs, k. Using the properties of the exponential process we obtain $\text{Prob}\{t \geq t_c\} = \exp(-t_c/\lambda)$ and $E[t \mid t \geq t_c] = \lambda + t_c$. Alternatively, we obtain the result by direct integration of $f(t) = \lambda^{-1}\exp(-t/\lambda)$. Thus, by substitution and simplification, the

$$\text{expected net rewards} = \gamma\alpha m p(\lambda + t_c - c)\exp(-t_c/\lambda) - k.$$

Differentiating the expected net rewards and setting the derivative to zero yields $t_c = c$. Finally, substitution yields Equation (1).

Equation (2). $E(u) = (1-p)\cdot 0 + p\cdot\text{Prob}\{t < t_c\}\cdot 0 + p\cdot\text{Prob}\{t \geq t_c\}E[u(\gamma\alpha m t - c)]$. Dropping terms that equal zero, substituting definitions, and using the properties of the exponential process, we obtain:

$$\begin{aligned} E[u] &= pe^{-t_c/\lambda}\int_{t_c}^{\infty}(1 - e^{-r\gamma m\alpha(t-c)})\lambda^{-1}e^{-(t-t_c)/\lambda}dt \\ &= pe^{-t_c/\lambda}\left[\int_0^{\infty}\lambda^{-1}e^{-x/\lambda}dx - e^{-rm\alpha\gamma(t_c-c)} \times \int_0^{\infty}e^{-(rm\alpha\gamma)x}\lambda^{-1}e^{-x/\lambda}dx\right]. \end{aligned} \tag{A1}$$

Recognizing the first integral as an integration over the range of a probability density function and the second integral as the Laplace transform of the exponential density, we obtain:

$$E[u] = pe^{-t_c/\lambda}\left[1 - \frac{e^{-rm\alpha\gamma(t_c-c)}}{1 + rm\alpha\gamma\lambda}\right]. \tag{A2}$$

Solving for the optimal cutoff ($t_c = c$) and substituting yields:

$$E[u] = rm\alpha\gamma\lambda pe^{-c/\lambda}\left[\frac{1}{1 + rm\alpha\gamma\lambda}\right]. \tag{A3}$$

For the constantly risk averse utility function, c.e. $= -(1/r)\log[1 - E(u)]$. If we substitute Equation (A3) into the expression for the c.e. and if we approximate $\log[1 - z] \approx z$ + second order terms, we get the result in the text. Because the utility function is constantly risk averse, we just subtract the certain costs, k. When the approximation does not hold, we use Equation (A3) directly.

Equation (3). The derivation of value to the firm follows that for the business unit manager except that $\gamma_F = 1$ and $\alpha = 1$. Equation (3) derives from the (risk neutral) condition that *expected value* $\geq$ *costs*.

Equation (4). With subsidies, the conditions for the business unit (manager) to select a project are $R\alpha m p\gamma\exp(-c/\lambda) \geq sk$. If $s = \alpha\gamma R$, then this is equivalent to Equation (3). If the business unit manager is risk neutral, then $R = 1$.

Tier 2: Development Programs to Match or Create Core Technological Competence

Equation (5). This equation is a definition. In the following, assume the tier 2 researchers are rewarded based on Equation (5), development metric $= \beta_v \tilde{m} v + \beta_e \tilde{e} - \beta_K K$.

Equation (6). Following the text we assume that the scope, $\tilde{m}$, results from M independent draws from a Bernoulli process with success probability q. Thus, the expected value and variance of the researcher's rewards are:

$$E[\text{rewards}] = \beta_v \Gamma^m M q v + \beta_e \Gamma^e e - \beta_K K - d(e),$$
$$\text{var}[\text{rewards}] = \beta_v^2 (\Gamma^m)^2 M q (1-q) v^2 + \beta_e^2 (\Gamma^e)^2 \sigma_e^2. \tag{A4}$$

We use the DeMoivre-Laplace Theorem (Drake 1967, p. 219) to represent the Bernoulli process outcomes with a normal approximation. For normally distributed outcomes and constantly risk averse utility functions, the c.e. = expected value − $(r/2)$(variance of outcomes). (The result is also approximate for other density functions.) For both results, see Keeney and Raiffa (1976, pp. 161, 202).

Figure 4. To illustrate the effect on program choice we, temporarily, ignore e by assuming that e and $d(e)$ do not vary by program. Thus, only β_v and β_K will affect program choice. See Footnote 5. For Figure 4, we normalize $e^* - w_o - d(e^*)$ to 0 as this does not change the basic concepts.

Figure 4a. To demonstrate the effect of discounting we assume risk neutrality and focus on E[rewards]. Because we are temporarily assuming risk neutrality, there are no risk costs to the researchers. Thus, the condition for positive profits for the firm simplifies to $v \geq K/(Mq)$. Furthermore, the choice of program by researchers will be based on only the β_v and β_K terms in E[rewards]. For the researchers who discount rewards, we rearrange E[rewards] to obtain the cutoff as $v \geq (\beta_K/\beta_v)K/(\Gamma^m Mq)$. The conditions for choosing program 2 over program 1 are then:

$$\text{Researcher: } v_2 \geq \frac{M_1 q_1 \Gamma_1^m}{M_2 q_2 \Gamma_2^m} v_1 + \frac{\beta_K}{\beta_v \Gamma_2^m} \frac{(K_2 - K_1)}{M_2 q_2},$$
$$\text{Firm: } v_2 \geq \frac{M_1 q_1}{M_2 q_2} v_1 + \frac{(K_2 - K_1)}{M_2 q_2}. \tag{A5}$$

Figure 4b. We now allow the researcher to be risk averse. If the firm could select the programs itself (and not rely on the implications of the development metrics) it would not have to pay risk costs. In this case, the firm's cutoff value would not change. However, if the firm relies on development metrics, the researcher's minimum c.e. condition becomes

$$\beta_v \Gamma^m M q v - (r/2) \beta_v^2 (\Gamma^m)^2 M q (1-q) v^2 \geq \beta_K K.$$

This quadratic equation will yield both a minimum cutoff (v too small) and a maximum cutoff (v too risky). That is,

$$v \geq [1 - \sqrt{1 - 2\beta_K K r (1-q)/Mq}]/[r \beta_v \Gamma^m (1-q)],$$
$$v \leq [1 + \sqrt{1 - 2\beta_K K r (1-q)/Mq}]/[r \beta_v \Gamma^m (1-q)]. \tag{A6}$$

Figure 4c. The conditions for choosing program 2 over program 1 become

$$\beta_v \Gamma_2^m M_2 q_2 v_2 - (r/2) \beta_v^2 \Gamma_2^{m_2} M_2 q_2 (1-q_2) v_2^2 - \beta_K K_2$$
$$\geq \beta_v \Gamma_1^m M_1 q_1 v_1 - (r/2) \beta_v^2 \Gamma_1^{m_2} M_1 q_1 (1-q_1) v_1^2 - \beta_K K_1.$$

This is a quadratic equation that will yield hyperbolic boundaries in (v_1, v_2)-space. For Figure 4c we have used the special conditions of $K_1 = K_2$ and $M_2 q_2 / M_1 q_1 = (1 - q_2)/(1 - q_1)$. These conditions reduce the boundaries to straight lines to demonstrate the regions more clearly. The intuitive reasoning for Figure 4c is that, for a given v_2, as v_1 gets very large, program 1 becomes less attractive due to risk. There will be regions where the researcher prefers a less risky program 2 over program 1, even though program 1 has a higher expected value.

Figure 4d. The effects from *both* figures 4b and 4c are plotted. The specific values used are $K_1 = K_2 = 2$, $M_1 q_1 = 10$, $M_2 q_2 = 8$, $\Gamma_1^m = 0.9$, $\Gamma_2^m = 0.6$, $\beta_v = \beta_K = 1$, and $r = 2$ (Figures 4b, c, d) or $r = 0$ (Figure 4a).

Equation (7). Because $\tilde{m}$ and $\tilde{e}$ are independently distributed, for a given choice of program, the firm can select β_e^* so that the researcher puts in the optimal effort e^* as given in Equation (8) below. To retain the researcher as an employee, the firm must pay at least market wages, w_o. However, to assure participation, the firm must reimburse its employee for the extranormal costs of effort, $d(e^*)$. In addition, the firm must reimburse the employee for any risk costs that its reward system induces the employee to undertake. Recognizing that the firm's market outcomes are $Mqv + e^*$ and that the firm (ultimately) bears the development cost, K, yields Equation (7). The subscript, j^*, indicates optimal program selection.

Equation (8). Following the arguments above, the certainty equivalent for a given e is given by:

$$\text{c.e.} = \beta_v \Gamma^e M q v - (r/2) \beta_v^2 (\Gamma^e)^2 M q (1-q) v^2$$
$$- \beta_K K + \beta_e \Gamma^e e - d(e) - r \beta_e^2 \sigma_e^2 / 2.$$

Since v is given, this reduces to c.e. = constant + $\beta_e \Gamma^e e - d(e)$. The tier 2 researcher will choose e such that $(\partial d/\partial e) = \beta_e \Gamma^e$. By the Implicit Function Theorem, this implies $(\partial e/\partial \beta_e) = \Gamma^e (\partial^2 d/\partial e^2)^{-1}$. In equilibrium, the firm must reimburse the tier 2 researcher for effort and risk costs, thus the firm will maximize $\{e - d(e) - r\beta_e^2 (\Gamma^e)^2 \sigma_e^2/2\}$. Recognizing that e is an implicit function of β_e, we solve this maximization problem to obtain Equation (8). We find the optimal efforts, e^*, by solving $(\partial d^*/\partial e) = [1 + r\sigma_e^2(\partial^2 d^*/\partial e^2)]^{-1}$.

Effort Indicators. Suppose that y and z are effort indicators, such as patents, publications, citations, or peer review, and suppose that y, z are jointly distributed as independent normal variates with variances, σ_y^2 and σ_z^2, respectively. Both have means of e. Holmstrom (1989) demonstrates that the optimal contract is linear in y and z. Using this fact, we derive the tier 2 researcher's optimal e for a given set of weights, a_y and a_z. This yields $(\partial d/\partial e) = a_y + a_z$. (Note that if $\Gamma^e = 1$ then the tier 2 researcher is rewarded now rather than later.) The Implicit Function Theorem yields $\partial e/\partial a_y = \partial e/\partial a_z = (\partial^2 d/\partial e^2)^{-1}$. The firm will then set wages to assure that the researcher participates, that is, the c.e. of the wages will at least equal the researcher's reservation wage. Because, in equilibrium, it must reimburse for effort and risk costs, the firm will maximize $\{e - d(e) - ra_y^2\sigma_y^2/2 - ra_z^2\sigma_z^2/2\}$. Recognizing that e is an implicit function, we solve this maximization problem to show that $a_y\sigma_y^2 = a_z\sigma_z^2$. This is the result quoted in the text.

Tier 1: Basic Research Explorations: The Role of Research Tourism

IMPLICATION 3. *When spillovers are possible, (a) the optimal number of explorations increases but (b) the optimal number of internal explorations might decrease.*

(a) We show $\partial n^*/\partial\mu > 0$. The firm wishes to maximize $V(n) - \kappa_i h - \mu\kappa_o h$ with $n = h + \mu h$. Hence, the firm maximizes $V(n) - [(\kappa_i + \mu\kappa_o)/(1 + \mu)]n$ which implies the optimality condition of $\partial V(n^*)/\partial n = (\kappa_i + \mu\kappa_o)/(1 + \mu)$. Implicit differentiation yields

$$\partial n^*/\partial\mu = [\partial^2 V(n^*)/\partial n^{*2}]^{-1}(\kappa_o - \kappa_i)/(1 + \mu)^2.$$

Thus, $\partial n^*/\partial\mu > 0$ because $V(n)$ is concave and $\kappa_o < \kappa_i$. (The firm prefers μ to be as large as possible because

$$\begin{aligned}&\partial\{V[h^* + \mu h^*] - \kappa_i h^* - \mu\kappa_o h^*\}/\partial\mu\\ &\quad= [\partial V(n^*)/\partial n - \kappa_o]h^*\\ &\quad= [(\kappa_i + \mu\kappa_o)/(1 + \mu) - \kappa_o]h^* > 0 \quad \text{for } \mu > 0.)\end{aligned}$$

(b) To prove the result we must only establish that an example exists such that internal explorations decrease. We establish existence with the example $V(n) = V_o \log(n + 1)$. Notice that $V(n = 0) = 0$. For this example we show that $\partial h^*/\partial\mu$ is ambiguous. In terms of h, the firm maximizes $\{V_o \log(h + \mu h + 1) - \kappa_i h - \kappa_o\mu h\}$. Differentiating and solving for h^* yields $h^* = V_o/(\kappa_i + \mu\kappa_o) - 1/(1 + \mu)$. For $h^* > 0$, this requires $V_o/(\kappa_i + \mu\kappa_o) > 1/(1 + \mu)$. Differentiating again we obtain

$$\partial h^*/\partial\mu = 1/(1 + \mu)^2 - (\kappa_o/V_o)(V_o^2/[\kappa_i + \mu\kappa_o]^2).$$

For $\kappa_o \to 0$, $\partial h^*/\partial\mu > 0$. For

$$\kappa_o \to \kappa_i, \quad \partial h^*/\partial\mu \to [1 - V_o/\kappa_i]/(1 + \mu)^2,$$

hence $\partial h^*/\partial\mu < 0$ whenever $V_o > \kappa_i$. This last condition is necessary for $n^* > 0$. (If n^* were not positive, there would be no need for tier 1.) □

IMPLICATION 4. *(a) If tier 1 researchers are evaluated on all ideas, new concepts, new technology, and new science, including those identified outside the firm, they will set $\mu_n^* = \bar{\mu}$ and invest in the optimal number of explorations for the firm. (b) If tier 1 researchers are evaluated on internal explorations only, they will adopt an NIH attitude by setting $\mu_h^* = 0$. They will work on more internal explorations and may develop fewer ideas, new concepts, new technology, and new science than would be optimal for the firm.*

(a) We first consider the case when tier 1 researchers are evaluated on $g_n(n) = V(n)$. Tier 1 researchers will select $\mu^o = \mu_n^*$ and $n = n^*$ to maximize

$$\{V(n) - \kappa_i n/(1 + \mu^o) - \kappa_o\mu n/(1 + \mu^o)\}.$$

Differentiating, we obtain:

$$\begin{aligned}&\partial\{V(n) - \kappa_i n/(1 + \mu^o) - \kappa_o\mu n/(1 + \mu^o)\}/\partial\mu^o\\ &\quad= n(\kappa_i - \kappa_o)/(1 + \mu^o)^2 > 0.\end{aligned}$$

Thus, it is optimal for the researcher to set $\mu_n^* = \bar{\mu}$. With $\mu^o = \bar{\mu}$ and $g_n(n) = V(n)$, the researcher's objectives match those of the firm.

(b) Let n^* and h^* be the optimal values that result from a metric of the form $g_n(n)$ and let n^o and h^o be the optimal values that result from a metric of the form $g_h(h)$. Following the text, we now consider the case when tier 1 researchers are evaluated on $g_h(h) = V[(1 + \bar{\mu})h]$ where $\bar{\mu}$ is announced by the firm as a parameter of the reward function. Tier 1 researchers select $\mu^o = \mu_h^*$ and $h = h^o$ to maximize $\{g_h(h) - \kappa_i h - \kappa_o\mu^o h\}$. Since

$$\partial\{g_h(h) - \kappa_i h - \kappa_o\mu^o h\}/\partial\mu^o < 0,$$

researchers will set $\mu_h^* = 0$. The revised choice of h^o is given by $\partial V(h^o)/\partial n = \kappa_i/(1 + \bar{\mu})$. The firm's optimal is given by $\partial V(n^*)/\partial n = (\kappa_i + \kappa_o)/(1 + \bar{\mu})$ where $n^* = (1 + \bar{\mu})h^*$. Hence,

$$\partial V[h^o(1 + \bar{\mu})]/\partial n < \partial V[(1 + \bar{\mu})h^*]/\partial n.$$

Since $V(\cdot)$ is concave, this implies that $(1 + \bar{\mu})h^o > (1 + \bar{\mu})h^*$, hence $h^o > h^*$. We establish the ambiguity of the comparison of n^o with n^* by using the example from Implication 3 to prove existence. Because the metric induces researchers to set $\mu_h^* = 0$, $n^o = h^o$, we compute $h^o = [V_o(1 + \bar{\mu}) - \kappa_i]/(\kappa_i + \kappa_i)$ and $n^* = [V_o(1 + \bar{\mu}) - \kappa_i - \kappa_o]/(\kappa_i + \kappa_o)$. As $\kappa_o \to \kappa_i$, $h^o > n^*$. As $\kappa_o \to 0$, $n^* > h^o$ by the condition that $h^* > 0$. □

Footnotes 9 and 10. *If $g_h(h) = V(h)$, then tier 1 researchers may work on more internal explorations and will develop fewer ideas than would be "optimal" for the firm.*

These footnotes cover the case where the firm rewards only on internal ideas, but anticipates that tier 1 researchers will adopt NIH and set $\mu_h^* = 0$. Under these conditions (and no risk aversion) the firm will select $g_h(h) = V(h)$. With $\mu_h^* = 0$, tier 1 researchers maximize $V(h) - \kappa_i h$, hence $\partial V(h^o)/\partial h = \kappa_i$ and $n^o = h^o$. The firm's optimal is given by $\partial V(n^*)/\partial n = (\kappa_i + \kappa_o)/(1 + \bar{\mu})$. Thus, $h^o = n^o < n^*$ because $V(\cdot)$ is concave and $\kappa_i > (\kappa_i + \kappa_o)/(1 + \bar{\mu})$. We establish the ambiguity of h^o vs. h^* with the example of Implication 3. We first compute $h^o = (V_o - \kappa_i)/\kappa_i$ and

$$h^* = [V_o(1 + \bar{\mu}) - \kappa_i - \kappa_o]/[(\kappa_i + \kappa_o)(1 + \bar{\mu})].$$

As $\kappa_o \to 0$, $h^* > h^o$ and as $\kappa_o \to \kappa_i$, $h^* < h^o$. □

References

Abt, R., M. Borja, M. M. Menke, J. P. Pezier. 1979. The dangerous quest for certainty in market forecasting. *Long Range Planning* **12** 2 (April) 52–62.

Acs, Z. J., D. B. Audretsch, M. P. Feldman. 1992. Real effects of academic research: Comment. *Amer. Econom. Rev.* **82** (March) 363–367.

Adler, P. S., D. W. McDonald, F. MacDonald. 1992. Strategic management of technical functions. *Sloan Management Rev.* **33** 2 (Winter) 19–37.

Allio, R. J., D. Sheehan. 1984. Allocating R&D research effectively. *Res. Management* (July–Aug.) 14–20.

Bachman, Paul W. 1972. The value of R&D in relation to company profits. *Res. Management* **15** (May) 58–63.

Bean, A. S. 1995. Why some R&D organizations are more productive than others. *Res. Tech. Management* **38** 1 (Jan–Feb) 25–29.

Bernstein, J. L., M. Ishaq Nadiri. 1989. Research and development and intra-industry spillovers: An empirical application of dynamic duality. *Rev. Econom. Studies* (April) 249–269.

Block, Z., O. A. Ornati. 1987. Compensating corporate venture managers. *J. Bus. Venturing* **2** 1 41–51.

Boblin, N. H., H. J. Vantrappen, A. E. Wechsler. 1994. The chief technology officer as an agent of change. *Prism* (Fourth Quarter) 75–85.

Boschi, Roberto A. A., H. U. Balthasar, M. M. Menke. 1979. Quantifying and forecasting research success. *Res. Management* **22** 5 (Sept.) 14–21.

Boulding, W., R. Morgan, R. Staelin. 1997. Pulling the plug to stop the new product drain. *J. Marketing Research* **34** 1 (February) 164–176.

Braunstein, D. M., M. C. Salsamendi. 1994. R&D planning at ARCO chemical, *Res. Tech. Management* **37** 5 (Sept–Oct), 33–37.

Brown, J. S. 1991. Research that reinvents the corporation, *Harvard Bus. Rev.* **69** 1 (Jan–Feb) 102–111.

Chester, A. N. 1994. Aligning technology with business strategy, *Res. Tech. Management* **37** 1 (Jan–Feb) 25–32.

Chester, A. N. 1995. Measurements and incentives for central research, *Res. Tech. Management* **38** 4 (July–Aug) 14–22.

Cohen, W. M., D. A. Levinthal. 1989. Innovation and learning: The two faces of R&D. *Econom. J.* **99** 39 569–596.

Cooper, R. G., E. J. Kleinschmidt. 1995. Benchmarking the firm's critical success factors in new product development. *J. Product Innovation Management* **12** 5 374–391.

Corcoran, E. 1994. The changing role of U.S. corporate research labs. *Res. Tech. Management* **37** 4 (July–Aug) 14–20.

David, H. A. 1970. *Order Statistics*. Wiley, New York.

Drake, A. W. 1967. *Fundamentals of Applied Probability Theory*. McGraw-Hill, New York.

Edwards, S. A., M. W. McCarrey. 1973. Measuring the performance of researchers. *Res. Management* **16** 1 (Jan) 34–41.

Erickson, G., R. Jacobson. 1992. Gaining comparative advantage through discretionary expenditures: The returns to R&D and advertising. *Management Sci.* **38** 9 (September) 1264–1279.

European Industrial Research Management Association (EIRMA) 1995. *Evaluation of R&D Projects*, Working Group Report No. 47.

Frohman, A. L. 1980. Managing the company's technological assets. *Res. Management* **23** 5 (May–June) 20–24.

Galloway, E. C. 1971. Evaluating R&D performance—keep it simple. *Res. Management* (March) 50–58.

Grabowski, H. G., J. Vernon. 1990. A new look at the returns and risks to pharmaceutical R&D. *Management Sci.* **36** 7 804–821.

Griffin, A., J. R. Hauser. 1996. The marketing/R&D interface. *J. Product Innovation Management*. **13** 3 (May).

Griliches, Z. 1990. Patent statistics as economic indicators: A survey. *J. Econom. Literature* **28** 4 1661–1707.

——. 1992. The search for R&D spillovers. *The Scandinavian J. Econom.* **94** Supplement, 29–50.

Gross, I. 1972. The creative aspects of advertising. *Sloan Management Rev.* **14** 1 (Fall) 83–109.

Gumbel, E. J. 1958. *Statistics of Extremes*. Columbia University Press, New York.

Hauser, J. R. 1996. Metrics to value R&D: An annotated bibliography. Working Paper, International Center for Research on the Management of Technology, MIT Sloan School, Cambridge, MA 02142 (March). Also available from the Marketing Science Institute.

——, D. I. Simester, B. Wernerfelt. 1994. Customer satisfaction incentives. *Marketing Sci.* **13** 4 (Fall) 327–350.

——, ——, ——. 1996. Internal customers and internal suppliers. *J. Marketing Res.* **33** 3 (August) 268–280.

Henderson, R., I. Cockburn. 1996. Scale, scope, and spillovers: The determinants of research productivity in drug discover. *Rand J. Econom.* **27** 1 (Spring) 32–59.

Hodder, J. E., H. E. Riggs. 1985. Pitfalls in evaluating risky projects. *Harvard Bus. Rev.* (Jan.–Feb.) 128–136.

Holmstrom, B. (1989), Agency costs and innovation, *J. Econom. Behavior and Organization* **12** 3 305–327.

Hultink, E. J., H. S. J. Robben. 1995. Measuring new product success: The difference that time perspective makes. *J. Product Innovation Management* **12** 5 392–405.

Irvine, J. 1988. *Evaluating Applied Research: Lessons from Japan*. Pinter Publishers, London.

Jaffe, A. B. 1986. Technological opportunity and spillovers of R&D: Evidence for firms patents, profits, and market value. *Amer. Econom. Rev.* **76** 5 (December) 984–1001.

——. 1989. Real effects of academic research. *Amer. Econom. Rev.* **79** 5 (December) 957–970.

Keeney, R. L., H. Raiffa. 1976. *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*. John Wiley & Sons, New York.

Koenig, M. E. D. 1983. A bibliometric analysis of pharmaceutical research. *Res. Policy* **12** 1 15–36.

Krause, I., J. Liu. 1993. Benchmarking R&D productivity: Research and development; case study. *Planning Rev.* **21** 1 (January) 16.

Krogh, L. C., J. H. Prager, D. P. Sorensen, J. D. Tomlinson. 1988. How 3M evaluates its R&D programs. *Res. Tech. Management* **31** 6 (Nov–Dec) 10–14.

Mansfield, E. 1980. Basic research and productivity increase in manufacturing. *Amer. Econom. Rev.* **70** 5 (December) 863–873.

——. 1981. Composition of R&D expenditures: Relationship to size of firm, concentration, and innovative output. *Rev. of Econom. and Stat.* **63** 4 (Nov) 610–615.

——. 1982. How economists see R&D. *Res. Tech. Management* **15** 4 (July) 23–29.

McGrath, M. E., M. N. Romeri. 1994. The R&D effectiveness index: A metric for product development performance. *J. Product Innovation Management* **11** 213–220.

Mechlin, G. F., D. Berg. 1980. Evaluating research—ROI is not enough. *Harvard Bus. Rev.* **59** (Sept–Oct) 93–99.

Miller, R. 1992. The influence of primary task on R&D laboratory evaluation: A comparative bibliometric analysis. *R&D Management* **22** 1 (January) 3.

Mitchell, G. R., W. F. Hamilton. 1988. Managing R&D as a strategic option. *Res. Tech. Management* **3** 3 (May–June) 15–22.

Negroponte, N. 1996. Where do new ideas come from. *Wired* (January) 204.

Nelson, R. 1959. The simple economics of basic scientific research. *J. Political Economy* **67** 297–306.

Pappas, R. A., D. S. Remer. 1985. Measuring R&D productivity. *Res. Management* **28** 3 (May–June) 15–22.

Patterson, W. 1983. Evaluating R&D performance at alcoa laboratories. *Res. Management* **26** 2 (March–April), 23–27.

Pien, H. 1997. Competitive advantage through successful management of R&D. Master's Thesis, Management of Technology Program, MIT, Cambridge, MA 02139.

Platt, J. 1964. Strong inference. *Science* **146** (October 16) 347–353.

Ransley, D. L., J. L. Rogers. 1994. A consensus on best R&D practices, *Res. Technology Management* (Mar–Apr) 19–26.

Reynolds, W. B. 1965. Research Evaluation. *Res. Management* (March) 117–125.

Rotemberg, J. J., G. Saloner. 1995. Overt interfunctional conflict (and its reduction through business strategy). *Rand J. Econom.* **26** 4 (Winter) 630–653.

Roussel, P. A., K. N. Saad, T. J. Erickson. 1991. *Managing the Link to Corporate Strategy: Third Generation R&D*, Harvard Business School Press, Boston, MA.

Schainblatt, A. H. 1982. How companies measure the productivity of engineers and scientists, *Res. Management* **25** 5 (May) 10–18.

Schmitt, R. W. 1987. R&D in a competitive era, *Res. Management* **30** (Jan–Feb) 15–19.

Sen, F., A. H. Rubenstein. 1989. External technology and In-house R&D's facilitative role, *J. Product Innovation Management* **6** 2 123–138.

Shapria, R., S. Globerson. 1983. An incentive plan for R&D workers. *Res. Management* (Sept–Oct) 17–20.

Stahl, M. J., J. Steger. 1977. Measuring innovation and productivity—A peer rating system. *Res. Management* (January).

Steele, L. W. 1987. Selling technology to your chief executive. *Res. Management* **30** 1 (Jan–Feb).

——. 1988. What we've learned: Selecting R&D programs and objectives. *Res. Tech. Management* (March–April) 1–36.

Stigler, G. 1961. The economics of information. *J. Political Econom.* **60** (June) 213–225.

Szakonyi, R. 1990. 101 tips for managing R&D more effectively—I, *Res. Tech. Management* **33** 4 (July–Aug) 31–36 and **33** 6 (Nov–Dec) 41–46.

Taylor, C. 1995. Digging for golden carrots: An analysis of research tournaments. *Amer. Econom. Rev.* **85** 4 (September) 872–890.

Tenner, A. R. 1991. Quality management beyond manufacturing. *Res. Tech. Management* **34** 5 (Sept–Oct) 27–32.

Tipping, J. W. 1993. Doing a lot more with a lot less. *Res. Tech. Management* **36** 5 (Sept–Oct) 13–14.

——, E. Zeffren, A. R. Fusfeld. 1995. Assessing the value of your technology. *Res. Tech. Management* **38** (5) 22–39.

Utterback, J. M. 1994. *Managing the Dynamics of Innovation*. Harvard Business School Press, Boston, MA.

Vest, C. M. 1995. Drift toward mediocrity in science. *The MIT Report* **23** 7 (Sept–Oct) 3–4.

von Hippel, E. 1988. *The Sources of Innovation*. Oxford University Press, New York.

Ward, M., D. Dranove. 1995. The vertical chain of R&D in the pharmaceutical industry. *Econom. Inquiry* **33** 1 (January) 1–18.

Whelen, J. M. 1976. Project profile reports measure R&D effectiveness. *Res. Management* **19** 5 (September) 14–16.

Zettelmeyer, F., J. R. Hauser. 1995. Metrics to value R&D groups, phase I: Qualitative interviews. Working Paper, International Center for Research on the Management of Technology, MIT Sloan School, Cambridge, MA.