

Metrics Thermostat

by

John R. Hauser

December 2000

John R. Hauser is the Kirin Professor of Marketing, Massachusetts Institute of Technology, Sloan School of Management, 38 Memorial Drive, E56-314, Cambridge, MA 02142, (617) 253-2929, (617) 258-7597 fax, jhauser@sloan.mit.edu.

This research was funded by the Center for Innovation in Product Development (CIPD) and the International Center for Research on the Management of Technology (ICRMOT), M.I.T. We wish to thank the managers and product development teams who worked with MIT to collect detailed metrics, covariates, and profit measures for the products that they launched. This article has benefited from presentations at the M.I.T. Marketing Group Seminar, the CIPD Thrust 3 Seminar, the NSF Site Review of CIPD, the ICRMOT International Seminar on Strategic Alliances and Metrics in Mishima Japan, the DuPont Speaker Series in Marketing at the University of Michigan, The Marketing Science Conference at Syracuse University, the Marketing Science Conference at the University of California at Los Angeles, the Boston Chapter of the Society of Concurrent Engineering at GTE in Waltham, MA, the Marketing Science Institute's Marketing Metrics Workshop, the Office of Naval Research, and the US Airforce's Lean Sustainment Initiative. Special thanks to Burt LaFountain who collected the data that is analyzed in the primary application in this article and performed many of the initial analyses. Thanks also to Arpita Majumder, who helped with the data analysis and who collected the data for an application to the automobile company, to Lcdr. Carl Frank, Lcdr. Keith Russell, 2nd Lt. Jeffrey Moffit, and Tom Kowalczyk who are applying the metrics thermostat to the US Navy and the US Air Force, to David Godes and Robert Zeithammer who gave me detailed feedback on details of this article, to Sandy Campbell, Conger Gable, Ken Dabrowski, Maurice Holmes, Gerry Katz, and Robert Klein who provided valuable managerial interpretations and who have suggested improvements and further applications.

Metrics Thermostat Author Biography

John R. Hauser is the Kirin Professor of Marketing and Head of the Marketing Group at M.I.T.'s Sloan School of Management where he teaches new product development, marketing management, competitive marketing strategy, and research methodology. He is the co-author of two textbooks, *Design and Marketing of New Products* and *Essentials of New Product Management*. He has published over sixty scientific papers. Among his awards include the Converse Award for contributions to the science of marketing, the MSI award for the most significant contribution to practice, three awards for the best paper of the year in the marketing sciences, and four additional awards as finalist. He has won an award from the Sloan School for outstanding teaching in the Master's program. His students have won major thesis awards including the Brooke's Prize, the Zennetos' Prize, the American Marketing Association dissertation award, and the INFORMS Bass Award. He has consulted for a variety of corporations on product development, sales forecasting, marketing research, voice of the customer, defensive strategy, and R&D management. He is one of the founders of Applied Marketing Sciences, Inc and has provided expert testimony regarding marketing research and sales forecasting. His current interests include metrics for R&D and for product development, customer satisfaction incentives, new methods for targeting customer needs, and the development of an Internet-based "virtual customer" that will provide rapid, usable information to product development teams..

Metrics Thermostat

Abstract

The explosion of information and information technology has led many firms to evolve a dispersed product development process with people and organizations spread throughout the world. To coordinate such dispersed processes a critical role of managers is to establish and foster a culture that implicitly rewards and motivates product development teams to perform against a set of strategic metrics such as customer satisfaction, time to market, defect reduction, or platform reuse. We focus on a practical method to fine-tune a firm's relative emphasis on these metrics. In particular, we seek to advise a firm how to modify their emphasis on each metric in order to improve profits. We use a thermostat analogy based on an adaptive control feedback mechanism in which we estimate the incremental improvements in priorities that will increase profits. Iterations seek to maximize profits even if the environment is changing.

In developing the metric thermostat we recognize that there are hundreds of detailed actions, such as the use of the house of quality and the use of robust design, among which the product development team must choose. We also recognize that team members will act in their own best interests to choose the actions that maximize their own implicit rewards as determined by the metrics. Management need not observe or dictate these detailed actions, but rather can control the process by establishing implicit weights on the metrics. The thermostat works by changing those implicit weights.

We define the problem, introduce the adaptive control mechanism, modify “agency” theory to deal with incremental changes about an operating point, and derive methods that are practical and robust in light of the data that firms have available. Our methods include statistical estimation and internal surveys. The mathematics identify the critical few parameters that need be determined and highlight how to estimate them. Both the measures and the estimation are illustrated in an initial application to a large office-equipment firm with \$20 billion in revenue. The metrics thermostat suggests that this firm has about the right emphasis on time-to-market, but has overshot on platform reuse and has lost its focus on customer satisfaction. We describe how the firm reacted to the recommendations and changed its organization.

This new product development vision is ... about people working in a completely new way in an environment where traditional barriers to remote communication and collaboration are essentially eliminated. It is about a major cultural reversal away from the era of exclusion, control, and co-location that product development managers worked so hard to build over the last 30 years.

*Maurice Holmes, former Chief Engineer, Xerox Corporation
Keynote Address, PDMA 1999 International Conference*

Managing a Dispersed Product Development Process with Metrics

We have spoken to senior managers at Ford, IBM, General Motors, ITT, Xerox, AOL, and the US Navy. They share Mr. Holmes belief that the explosion in information and information technology including, but not limited to, the worldwide web, is transforming the way product development teams are working. With barriers to remote communication and collaboration reduced, there is a cultural shift to less centralized control. Their new vision is one of dispersed, self-directed, more-autonomous teams which are coordinated through common goals. We call those goals, metrics. A critical role of central management is to establish and foster a culture that motivates and rewards product development teams to allocate the right amount of effort toward achieving those metrics. For example, Maurice Holmes, in his role as Chief Engineer of Xerox, implemented a successful time-to-market (TTM) process that reduced cycle time by a factor of 2.5. He did this by encouraging the use of formal processes, by training, and most importantly, by ensuring that the Xerox product development teams knew that TTM was a critical goal. We spoke to many Xerox product development professionals during this period and it was clear that they knew and understood that activities which improved TTM would be rewarded well.

But TTM is not the only metric that affects product development success nor is it clear that TTM should be pursued to the exclusion of all other metrics. For example, Menezes [40] published a case documenting Xerox's 1980s shift from a single goal of ROI to a focus on customer satisfaction (CS). Menezes reports a major change in culture at Xerox and a successful transformation of Xerox's ability to compete in international markets. While new processes such as supplier integration, inter-functional coordination, product differentiation, brand-image management, service management, standards compliance, and the use of "rigorous" methods such as conjoint analysis, the house of quality, robust-design methods, design structure matrices, and platform reuse matrices might improve both TTM and CS, some teams might choose to sacrifice customer satisfaction in their rush to get a product to market. For example, they might save a month by substituting a single focus group for a more rigorous and complete analysis of customer needs. This month might cost them dearly in their

ability to satisfy customers. There are similar tradeoffs made every day among metrics such as time to market, customer satisfaction, defect reduction, platform reuse, core-competence alignment, reduced product complexity, enhanced variety, service quality, and the more than eighty variables that have been identified in the product development literature as affecting success [1, 2, 5, 7, 9, 10, 11, 12, 13, 19, 21, 22, 23, 24, 26, 32, 33, 36, 37, 42, 43, 45, 46, 49, 51, 53, 58, 59, 60, 61].

Selecting the right metrics, establishing a culture that rewards those metrics, and tuning that culture to place the right relative emphasis on each metric are critical tasks in managing a dispersed product development process. And these tasks are not easy! The thirty articles referenced in the previous paragraph are but a sampling of the rich literature addressing metrics. Selecting metrics and establishing a culture are extremely important and deserve the attention they have gotten in the literature. We hope to contribute to that literature by focusing on the third task, tuning the culture.

In the four applications that we have studied to date, we have found that management had invested considerable resources to select the metrics by which to manage, usually in formal committees and special studies often lasting many months. Management had also invested considerable resources in training and incentives, often over \$100 million per metric, to establish the metric-based culture. Indeed, the CEO of a company with over \$4 billion in revenues told us that establishing the right culture is one of his most important tasks. He regularly attends internal seminars on robust processes in order to lead by example and show that the robustness metric is important to his firm. But these same managers are concerned with whether they are placing the right emphasis on each metric. They are well-aware that metrics such as TTM and CS might conflict. They want to fine-tune their culture so that each product development team acting autonomously (and in its own best interests) takes the actions and makes the decisions with respect to these metrics in a manner that maximizes the overall long-term profit of the firm.

These managers want a practical method. Many are aware of the scientific literature that studies the antecedents and consequences of metrics in multi-firm studies. But to fine-tune the culture of their firm, these managers need a method that adjusts priorities based on measures of their firm. This need presents an empirical challenge. The typical firm, even if it has over \$20 billion in revenues, does not launch hundreds of products each year. Furthermore, in some firms, cultures (and the implicit weights on metrics) vary by division or even within a division. For example, one division might be a defense contractor serving primarily government customers while another division of the same firm might serve consumer markets with millions of customers. In those firms, our method must be able to select the right emphasis for each division or subset of projects.

In this article we explore one practical method to fine-tune the emphasis that a given firm places on its product development metrics. We assume that the firm has already selected the metrics by which to manage and, if given directions to increase or decrease emphasis on a metric, they have the means to do so. We take as a constraint that we will have limited data – often only twenty or fewer products that have been launched by the firm (division) within the last five years. We develop a model that is robust and conservative. If the data are too noisy the model should not falsely recommend change – instead management should rely on the extant methods they now use. In the following sections, we describe the basic method, derive the underlying mathematics, and apply the method to a large technological firm.

A Metrics Thermostat

Consider the thermostat in the room in which you are reading this article. Sometime, perhaps many years ago, an HVAC engineer designed a heating and cooling system for your building. He or she drew on a rich scientific base and on years of experience to study the characteristics of the room, the building, the climate, the likely occupants, and many other variables to select heating, cooling, and ventilating equipment. The system was a good engineering solution for your building (we hope), but not for every minute of every day. To fine-tune the system the engineer placed thermostats in each room. Each thermostat simply reads the temperature in the room, compares it to a target (optimal) temperature, and sends a signal to increase or decrease heat (or cooling) to the room. If the sensitivity is designed well (most thermostats allow a variation of a few degrees before “requesting” changes), the room will track the optimal temperature even when the external input (such as the outside temperature and sun load) varies dramatically.

We draw on this analogy with the system below the dotted line in Figure 1. We enter the process after management has established a product development process and selected the metrics by which to manage (boxes above the dotted line). This article focuses below the dotted line as management tunes the culture that sets the implicit rewards and incentives. For example, at Xerox members of teams often had a better chance of being promoted or getting assigned to the most interesting and prestigious projects if they did well on TTM. Conceptually, the metrics thermostat provides the feedback in the system necessary to fine-tune that culture.

The “thermostat” measures profits, both short-term and projected long-term, and compares them to measures of the metrics. Based on these comparisons and on other measures of the organization (described later) the thermostat makes suggestions to increase or decrease the emphasis on each metric. After management adjusts the culture, another measurement-and-adjustment cycle begins.

For example, if the thermostat suggests an increase in emphasis on customer satisfaction, management might encourage the greater use of voice-of-the-customer methods or might institute a formal customer-satisfaction incentive system. If the thermostat suggests that the culture has overshot on customer satisfaction – perhaps because the teams are spending too much costly effort studying customers and the measured benefits do not justify the costs – then, management might respond by lowering its relative emphasis on customer satisfaction. If the thermostat has the right sensitivity, these cycles of adjustment should maximize profits.

The idea is simple, but to implement a thermostat we must address three issues. First, our system must recognize that while management rewards metrics such as customer satisfaction, the team takes detailed actions such as the use of the house of quality or the use of a particular software solution. Second, unlike an HVAC system, members of a product development team have their own personal goals and ambitions. These may or may not be compatible with the firm’s goals of profit maximization. Our system must take these “agency” issues into account in a form that matches the measures we can obtain. Third, while an HVAC thermostat responds to a single metric, temperature, and turns the system off or on, the product development teams responds to many metrics simultaneously and we must adjust the relative emphasis among the metrics. We address each of these issues in turn beginning with the mapping from actions to metrics.

Adaptive Control of the System: Actions → Metrics → Priorities

The following example is designed to be simple so that the framework is clear. We later demonstrate adaptive control with a more complex empirical application.

There are literally hundreds of types of actions that new-product development teams can take to affect outcomes. These actions include process steps such as conjoint analysis, perceptual mapping, platform reuse methods, design structure matrices, the house of quality, and Taguchi methods. Actions also include investigating new composite materials, new electronic circuits, or improved customer-service procedures. We label actions as a_k for $k = 1$ to K . For complex products, K is a large number. However, for the purposes of this illustration we consider only two actions, a_1 , the use of the house of quality (HOQ) and, a_2 , the application of reuse matrices (RUM) to reduce costs and speed products to market [20, 27].

There is likely a profit-maximizing allocation of effort among the actions. For example, platform architecture methods such as reuse matrices might require careful planning but could bring about efficiencies that reduce the average time and cost of the products in a family. But, if reuse matrices are overused they might lead to less innovation and a lesser ability to satisfy a diverse customer

base. Similarly, it is hard to imagine new product success without some customer input, but formal house of quality methods sometimes take too long and cost too much. Putting this qualitative experience together, we expect that a plot of profit versus the use of HOQ and RUM might look something like Figure 2a.

Figure 2a is simple to visualize. If management could do sufficient experiments to obtain this plot and if they could monitor the product development teams perfectly, then they would simply dictate the optimal actions. But management can not know the details of this plot and can not simply dictate actions. Furthermore, practical problems involve hundreds of actions taken by team members with detailed knowledge of the methods and the problems being addressed. It would be very hard indeed to dictate these actions to the team – especially under dispersed product development processes. Instead of dictating actions, management selects strategic metrics, m_i for $i = 1$ to n . Management chooses metrics that are (1) correlates of profit, (2) measurable, and (3) affected by the team's actions. In real applications we expect that $n \ll K$, but for this section we illustrate the method with $n=K$ by selecting two metrics, m_1 , customer satisfaction (CS), and m_2 , time to market (TTM).

For our illustration we have chosen “soft” metrics as recommended by the performance measurement interpretation of Baker [3] and Gibbons [18] that expands on ideas of Holmstrom and Milgrom [31]. Unlike the classical “agency” interpretation (e.g., Holmstrom [30]), soft metrics need not be noisy indicators of profit. Instead, we seek metrics that “induce the agent to do the right thing at the right time [18, p. 10].” This subtle, but important, change in interpretation enables the firm to choose metrics, such as CS or TTM, that impose less risk and fewer time delays than would be imposed by direct, but noisy measures of incremental profit. This is particularly important in product development where such metrics are measured well in advance (and with less noise) than long-term profits. The latter often occur many years after a team has been disbanded. (For those readers interested in a formal “repeated games” motivation of soft metrics see Gibbons [18].)

In theory, there is some mapping of actions into metrics that is understood by the product development teams. That is, teams know that every set of actions will produce some likely levels of the metrics. Indeed, when $n \ll K$, teams may know many alternative actions that produce the same set of metrics. For example, the team might compensate for less effort on the house of quality with more product testing. Given that the team is rewarded implicitly on the metrics, it is natural to assume they will strive to choose those actions that achieve the metrics at the least cost to themselves. This is the natural result of the trend toward a dispersed product development process that empowers the team to make detailed tradeoffs [5, 54]. Such team self interest implies that we can associate actions in action space with levels of metrics in metrics-space.¹

In general, the resulting mapping is non-linear, but we illustrate the concept with linear functions, $CS = \frac{3}{7} HOQ - \frac{3}{7} RUM$, $TTM = \frac{1}{7} HOQ + \frac{6}{7} RUM$. (In general, less TTM is perceived as better and rewarded accordingly. Thus, our illustrative metric, m_2 , measures performance on TTM such that a higher number on the metric is “better.”) In our notation, these equations are rewritten as:

$$(1) \quad m_1 = \frac{3}{7} a_1 - \frac{3}{7} a_2 \quad m_2 = \frac{1}{7} a_1 + \frac{6}{7} a_2$$

Equation 1 suggests that efforts to employ reuse matrices increase performance on time to market, but decrease customer satisfaction. On the other hand, house of quality efforts increase customer satisfaction and performance on time to market. These directional impacts are illustrative only. We address actual empirical relationships in the application.

Based on Equation 1, we transform Figure 2a from action-space to the metrics-space in Figure 2b.² The surface in Figure 2b is more complex than the surface in Figure 2a because the underlying equation now includes interactions (CS x TTM) that were not present before. This is a direct result of the tradeoffs in Equation 1. The small, heavy circle (●) in Figure 2b indicates the region in metrics space in which the organization is now operating. We call this the current operating point. In other words, the current culture, incentives, and leadership have encouraged teams to take the actions that lead to this set of metrics. In Figure 2b, the firm is not operating for maximum profit. It can improve. Our goal is to help it improve.

If the target firm, or a subset (division) of that firm, is sufficiently homogeneous, then all teams in that division will be operating in the neighborhood of the heavy circle. However, we do expect some variation. If there is sufficient variation, but not too much, then we can approximate the non-linear curve with the hyperplane in Figure 2c. (We formalize this later.) If we estimate how changes in the metrics affect profit in the hyperplane, then we could change the priorities on the metrics to encourage the teams to take actions that move the metrics in the direction of the vector in Figure 2c. To determine this hyperplane we need only estimate how profits change with respect to CS and TTM in the neighborhood of the operating point ($\partial\pi / \partial m_1$ and $\partial\pi / \partial m_2$). We do not have to determine the entire non-linear surface. If we make a small improvement, repeat the measurement, and continue to make improvements, then, under the right conditions, we will reach the optimal priorities on the metrics.

This measurement and adjustment system is called adaptive control. Adaptive control has a long history in marketing and has proven to be remarkably robust, even if the profit curve is changing over time. For example, Little [38] used adaptive control to optimize advertising spending and

showed that the system is robust, converges rapidly, and tracks optimality when the underlying parameters of the system change.³

Figure 2c is based on levels of the metrics, but management does not affect the metrics directly. They control the emphasis on the metrics through the implicit reward system inherent in the firm's culture. We model this as control of the weights, w_i 's, that summarize the strategic emphasis placed on the metrics, m_i 's. To effect control we must identify how to change the weights applied to the metrics (Figure 2d) because management cannot change the metrics directly (Figure 2c). This is known as an agency problem because management develops products through its "agents," the product development teams. We address the details of this agency problem in the next section. We find this easier to do in a more general notation.

In this article we have chosen to derive the formal equations rigorously so that they might be critiqued, extended, and improved by other researchers. However, the result of these derivations is a relatively simple formula that is conceptually easy to apply in real situations. Readers who are more interested in the practical application of the metrics thermostat than the formal derivations are encouraged to skip to the section, "Non-Mathematical Summary."

How Incremental Changes in Priorities Influence the Team to take Actions that Lead to Profitable Incremental Changes in the Metrics ("Agency" Issues)

In this section we focus on a single product development team (henceforth "team") taking actions for a single product development project (henceforth "PD project"). We expand to multiple teams and projects in the next section. We assume that management has chosen metrics that depend upon the team's actions (a_k 's) and has induced an implicit reward scheme. We assume the team maximizes its own rewards. The agency problem is to improve the relative priorities on these metrics such that the team's actions turn out to be in the best interests of the firm as illustrated by Figure 3 [4]. Figure 3 says that we want to encourage the team to take actions and make decisions (HOQ and RUM). These, in turn, affect soft metrics (TTM and CS). The actions also affect outcomes (short- and long-term profit). Because the firm rewards the team implicitly for performance on metrics (TTM and CS), the team chooses the actions (HOQ and RUM) to maximize its implicit rewards. If the firm's culture implies the right implicit rewards, the team's actions (HOQ and RUM) will also maximize the firm's profits. Note that the firm does not make money directly from the metrics (TTM or CS). Rather, the firm is profitable because the team takes the right actions (HOQ and RUM) to achieve the right levels of the metrics (TTM and CS).

Formally, the reward system induces the team to take a series of actions that produce levels of the metrics. Because the team is empowered and self-managed it chooses the actions, a_k 's, in its own

best interests. There is a cost, $c(a_1, a_2, \dots, a_K)$, to these actions. The team bears this cost and the details of the cost are not observable by management.

We introduce an intermediate construct to maintain the notation of agency theory. This construct makes the theory easier to present and understand but does not change the basic interpretations. The construct decomposes overall effort into its “coordinates” much as an engineer decomposes forces into their x-, y-, and z-components to make analysis easier. In this decomposition we assign an unobservable effort, e_i^a , to each metric, m_i . Mathematically, the e_i^a ’s summarize the teams’ actions (hence the a superscript) as if the team were allocating effort (through the actions) independently to each metric. We choose the e_i^a ’s such that each metric, m_i , depends only on one component of the effort decomposition, e_i^a . This decomposition is for mathematical convenience only – the real decisions and the real costs are with respect to the actions. We denote the efforts necessary for the current operating point as e_i^o and any incremental efforts to change the operating point as e_i . With this construct we transform the cost function: $c(a_1, a_2, \dots, a_K) \rightarrow c(e_1^a, e_2^a, \dots, e_n^a) = c(e_1^o + e_1, e_n^o + e_2, \dots, e_n^o + e_n)$. Without loss of generality, we write these costs more simply as $c^o(e_1, e_2, \dots, e_n)$.

Management’s measures of the metrics are imperfect – for example, a survey of customer satisfaction is, at best, an estimate of true customer satisfaction. Thus, the measure of each metric, \tilde{m}_i , after the team changes its actions, is a noisy measure that depends on the changes in the team’s efforts. That is, $\tilde{m}_i = m_i(e_i^a) + error_i$, where $error_i$ is a zero-mean normal random variable with variance σ_i^2 . We recognize that $m_i(e_i^a) = m_i(e_i^o + e_i)$. When the notation is clear we will write this as m_i . The current operating point is $m_i(e_i^o)$, or more simply, m_i^o . There is no need to model any error in the current operating point as both management and the team know the current measure and the team knows the actions it takes to achieve this operating point.

When applying adaptive control, we are operating in the tangent hyperplane of metrics-space, thus we use a Taylor’s series approximation of the optimal reward system about the current operating point. We define $w_i \equiv \partial rewards / \partial m_i$. The expansion is exact in an ϵ -neighborhood and a robust approximation close to the operating point. Thus, the team’s rewards can be written as follows.

$$(2) \text{ rewards}(\tilde{m}_1, \tilde{m}_2, \dots, \tilde{m}_n) \approx \text{rewards}(m_1^o, m_2^o, \dots, m_n^o) + w_1(\tilde{m}_1 - m_1^o) + w_2(\tilde{m}_2 - m_2^o) + \dots + w_n(\tilde{m}_n - m_n^o)$$

Collecting all constants as w_o , we rewrite Equation 2 as Equation 2’.

$$(2') \text{ rewards} \approx w_o + w_1\tilde{m}_1 + w_2\tilde{m}_2 + \dots + w_n\tilde{m}_n$$

Note that in Equation 2', the weights, w_i 's, represent incremental changes in the reward system that induced the team to allocate the efforts, $e_1^o, e_2^o, \dots, e_n^o$, that achieved $m_1^o, m_2^o, \dots, m_n^o$. Furthermore, we recognize that the rewards need not be, and rarely are, explicit monetary rewards. Rather they are often non-monetary incentives that the team values and that are costly for the firm to provide. Indeed, the firm in our example is sponsoring research to understand and quantify these non-monetary rewards. In preliminary research, Chan [8] applied Internet conjoint analysis to measure tradeoffs among non-monetary incentives.

If the measurement errors are uncorrelated and the team is constantly risk averse, then the team, acting in its own best interests, will maximize the following certainty equivalent (*c.e.*) where r is a risk aversion constant that quantifies the team's risk preference.⁴

$$(3) \quad c.e. \approx w_0 + w_1 m_1 + w_2 m_2 + w_n m_n - c^o(e_1, e_2, \dots, e_n) - \frac{1}{2} r w_1^2 \sigma_1^2 - \frac{1}{2} r w_2^2 \sigma_2^2 - \dots - \frac{1}{2} r w_n^2 \sigma_n^2$$

The actions, a_k 's, of the team lead to incremental profit which we write (using our "effort" construct) as $\pi(e_1^a, e_2^a, \dots, e_n^a)$. In the tangent hyperplane we use a Taylor's series expansion to obtain Equation 4 where $\partial / \partial e_i^o$ is shorthand for $\partial / \partial e_i^a$ evaluated at e_i^o .

$$(4) \quad \pi(e_1^a, e_2^a, \dots, e_n^a) \approx \pi(e_1^o, e_2^o, \dots, e_n^o) + \sum_{i=1}^n \frac{\partial \pi}{\partial e_i^o} (e_i^a - e_i^o) \equiv \pi^o + \sum_{i=1}^n \frac{\partial \pi}{\partial e_i^o} e_i$$

The firm can affect its net profits by selecting the constant, w_o , and the incremental changes in the weights, w_1, w_2, \dots, w_n . After paying these wages and bonuses, the firm's net profits in the neighborhood of the hyperplane are:

$$(5) \quad net \ profit \approx \pi^o + \sum_{i=1}^n \frac{\partial \pi}{\partial e_i^o} e_i - w_o - w_1 \tilde{m}_1 - w_2 \tilde{m}_2 - \dots - w_n \tilde{m}_n$$

Finally, we recognize that the firm will try to keep wages as low as feasible, hence it will choose w_o to maintain wages only as high as is necessary to prevent the team members from leaving the firm. That is, it will select $rewards \geq W_o$ where W_o represents the wages the team members could earn elsewhere after taking switching costs into account. This implies that the firm's formal maximization problem (for incremental changes) can be written as follows.

$$(6) \quad \max \ net \ profit \approx \pi^o + \sum_{i=1}^n \frac{\partial \pi}{\partial e_i^o} e_i - W_o - c^o(e_1, e_2, \dots, e_n) - \frac{1}{2} r w_1^2 \sigma_1^2 - \dots - \frac{1}{2} r w_n^2 \sigma_n^2$$

In order to solve the firm's maximization problem in Equation 6 (treating \approx as equality), we first solve the team's maximization problem, Equation 3. We again use a Taylor's expansion in the tangent hyperplane to obtain:

$$(7) \quad m_i(e_i^o + e_i) \approx m_i^o + \frac{\partial m_i}{\partial e_i^o} (e_i^a - e_i^o) \equiv m_i^o + \frac{\partial m_i}{\partial e_i^o} e_i$$

We now use standard agency theory to solve the firm's maximization problem. That is, we maximize Equation 3 to determine the optimal incremental efforts (e_i^* 's) that summarize the actions (a_k^* 's) the team will take to maximize their expected utility. We substitute these implied efforts in Equation 6 and derive the optimal incremental change in priorities (w_i^* 's) based on the properties of the profit, cost, and metrics functions. For a very readable review, see Gibbons [18]. For each metric, the optimal weight is given by the following equation.

$$(8) \quad w_i^* \approx \frac{\left(\frac{\partial \pi}{\partial e_i^o} \right) / \left(\frac{\partial m_i}{\partial e_i^o} \right)}{\left[1 + \left(r \frac{\partial^2 c^o}{\partial e_i^{o2}} \right) \left\{ \sigma_i / \left(\frac{\partial m_i}{\partial e_i^o} \right) \right\}^2 \right]}$$

Equation 8 has qualitative interpretations with useful engineering analogies. Product development engineers might call the term in the numerator, “leverage,” because it represents the ratio of the marginal change in incremental profits relative to the marginal change in the measure of the metric. For example, if the unobservable efforts move “customer satisfaction” one notch while moving incremental profit by five notches, then we say the customer satisfaction metric has a leverage of five.

The inverse of the right-most bracketed ratio in the denominator is similar to the engineering concept of signal-to-noise ratio (SNR). The partial derivative, $\partial m_i / \partial e_i^o$, defines the scale of the error-less signal from the metric while the standard deviation, σ_i , indicates the magnitude of the error (noise). If we were to rescale the metric, we would rescale both the partial derivative and the standard deviation by the same amount; hence the SNR is a dimensionless quantity. Equation 8 implies that an improved SNR will make w_i^* higher. The ideal metric would have both high leverage and high SNR, but real metrics require the firm to make tradeoffs. “Soft” metrics, such as customer satisfaction, might have higher leverage than “hard” metrics, such as the number of defects reported. The tradeoff is that the soft metrics will have lower SNRs. It becomes an empirical question as to whether the enhanced leverage justifies the degradation in SNR.

The remaining terms in Equation 8 quantify the team's desires to avoid risk, r , and effort, $\partial^2 c^o / \partial e_i^{o2}$. Because these terms are difficult to measure directly, we address next how to infer them

from other measures. (Eliashberg and Hauser [15] derive methods to estimate r directly, but measuring $\partial^2 c^o / \partial e_i^{o2}$ is much more difficult.)

Practical Measurement for Multiple PD Teams

In the previous section we modified the standard agency derivation to address the incremental (directional) changes necessary to implement adaptive control. We now use this mathematical machinery to address the managerial problem which requires practical measurement that applies to multiple teams. We address multiple PD teams because, practically, top management cannot set a different culture, incentive system, and leadership style for each team. The implicit weights set by the culture, incentives, and leadership are the same for a group of teams in a division or subset of the firm. Specifically, we let j index the projects undertaken by the teams that share the same culture, where $j=1$ to J . We add this index to actions, efforts, metrics, and profit. In addition, we recognize that each project may be affected by covariates and that incremental profit may be difficult to measure exactly. We model these errors by letting profit be a random variable such that $\tilde{\pi}_j = \pi_j + error_j^\pi$ where $error_j^\pi$ is a normal random variate. This error does not affect the derivation of Equation 8.

We simplify notation by using capital letters to represent the partial derivatives in Equation 8. Further, we drop all constants that do not affect the optimizations of Equations 3 and 6. Because our focus is to provide management with recommendations with respect to the weights, these derivatives are sufficient to provide the necessary approximations in the hyperplanes of Figures 2c and 2d.⁵

Let $M_{ij} \equiv \partial m_{ij} / \partial e_{ij}^o$, $\Pi_{ij} \equiv \partial \pi_j / \partial e_{ij}^o$, and $C_{ij}^2 \equiv \partial^2 c_j^o / \partial e_{ij}^{o2}$. Using these definitions we rewrite Equation 8 as:

$$(8') \quad w_{ij}^* = \frac{\Pi_{ij} / M_{ij}}{1 + r C_{ij}^2 \sigma_i^2 / M_{ij}^2}$$

Using Equation 3 we determine the corresponding optimal efforts: $e_{ij}^* = w_{ij}^* M_{ij} / C_{ij}^2$. Using Equations 2' and 7 we recognize that the expected incremental rewards to team j for taking the actions that affect metric i are a constant plus $w_{ij}^* M_{ij} e_{ij}^*$ while the certainty equivalent of those incremental rewards are $w_{ij}^* M_{ij} e_{ij}^* - \frac{1}{2} r (w_{ij}^*)^2 \sigma_i^2$. (Here we define the expected rewards and certainty equivalent as that obtained before subtracting unobserved costs. This is consistent with our empirical measurement.) Substituting in Equation 8' and simplifying gives a simple expression for the denominator.

$$(9) \quad w_{ij}^* = \frac{\Pi_{ij} / M_{ij}}{1 + 2 \left[\frac{E[\text{rewards}_{ij}^o] - c.e._{ij}^o}{E[\text{rewards}_{ij}^o]} \right]}$$

The bracketed term in the denominator is now a measurable quantity that we have come to call the *risk discount factor (RDF)*. For a given set of priorities, it is the amount by which the team will discount the real, risky rewards relative to a situation where the rewards can be guaranteed. We have pretested a number of measures of RDF and have found that team members understand the concept and provide consistent answers that they feel represent RDF. Table 1 reproduces the wording we used in our application. RDF is a measure of the current state of the organization which represents the net effect of risk aversion, effort aversion, the signal-to-noise ratio, and the current reward system. It applies in the tangent hyperplane at the current organizational operating point. It is remeasured for each iteration of adaptive control. (With experience, we hope future researchers will improve the questions in Table 1.)

To estimate the numerator of Equation 9 in the tangent hyperplane, we substitute the expression for e_{ij}^* into Equations 4 and 7 and collect terms to obtain:

$$(10) \quad \tilde{\pi}_j = \pi_j^o + \sum_{i=1}^n \frac{\Pi_{ij}}{M_{ij}} (\tilde{m}_{ij} - m_{ij}^o - error_{ij}) + error_j^\pi$$

Recognizing that π_j^o does not depend on the incremental changes in metrics we estimate it based on covariates outside the team's control such as the availability of resources, core competence alignment, size of strategic opportunity, and fit with corporate image. If we call these covariates, v_j^g , for $g=1$ to G , then profit at the organizational operating point can be approximated by a linear combination of the covariates, that is, $\pi_j^o = \sum_g \mu_g v_j^g + \text{constant} + error_j^g$, where μ_g is the weight on the j^{th} covariate. Thus,

$$(11) \quad \tilde{\pi}_j = \text{constant} - \sum_{i=1}^n \frac{\Pi_{ij}}{M_{ij}} m_{ij}^o + \sum_{i=1}^n \frac{\Pi_{ij}}{M_{ij}} \tilde{m}_{ij} + \sum_{g=1}^G \mu_g v_j^g + \sum_{i=1}^n (error_j^\pi + error_j^o - \frac{\Pi_{ij}}{M_{ij}} error_{ij})$$

To proceed further empirically, we must decide which parameters we can assume to be homogeneous within the hyperplane. For our applications we chose one set of assumptions that management felt was reasonable. Others may wish to extend the model with other assumptions and/or estimation procedures. In our case, management felt that the ratio, Π_{ij}/M_{ij} , would not vary dramatically across the PD projects in the division and that the operating point, m_{ij}^o , would apply across projects. Based on these approximations, Equation 11 becomes a simple multiple regression equation in which observed profits are regressed on the metrics and covariates. Thus, we estimate Π_i/M_i by $\hat{\lambda}_i$ in the following regression equation where $error_j'$ is itself a zero-mean, normal random variable.⁶

$$(12) \quad \tilde{\pi}_j = constant' + \sum_{i=1}^n \lambda_i \tilde{m}_{ij} + \sum_{g=1}^G \mu_g v_j^g + error_j'$$

If RDF_i is measured (and relatively homogeneous) within the hyperplane, then we have a simple, practical expression to update the weights for the division. (We use the superscript, d , to indicate that this is an estimate for a subset [division] of the firm that shares the same culture and we use $\hat{\cdot}$ to indicate that quantities are estimated.)

$$(13) \quad \hat{w}_i^d = \frac{\hat{\lambda}_i}{1 + 2RDF_i} \equiv \frac{(\Pi_i/M_i)^\wedge}{1 + 2 * (risk\ discount\ factor)}$$

These estimates update priorities in each iteration of adaptive control.

Non-Mathematical Summary

The mathematics, although a bit complicated, give a simple procedure. Equation 13 says that the amount by which we change priorities on the i^{th} metric, say customer satisfaction, is summarized by two measurable quantities. The first is the incremental change in profit that is measured when teams take actions that lead to incremental changes in metric i . To measure this quantity we first observe profits and metrics for the division's products and then use multiple linear regression to estimate how profit changes as the metrics and covariates change. The regression coefficient for metric i is the "leverage" of the metric. Equation 13 also says that we must modify this estimated leverage to take into account the team's self interest. We use agency theory to design a survey (Table 1) to measure the "risk discount factor" for each metric and we reduce the regression coefficient by dividing it by $(1 + 2RDF_i)$. This correction assures that the more risky the metric the less we change its priority. The resulting estimate (\hat{w}_i^d) tell us how much to change the culture's emphasis on metric i . This implies that we implement adaptive control (Figure 1) with the following 7-step procedure.

1. Identify a set of product development projects that follow approximately the same culture, that is, the teams are rewarded implicitly on the same metrics with approximately the same priorities.
2. Identify the metrics by which the firm is managed. That is, we select the metrics upon which the product development teams perceived they are implicitly rewarded.
3. Use the firm's documentation to obtain measures of the metrics, the covariates, and short- and long-term profit for each project in the last Y years (typically $Y=5$).
4. Use multiple regression on this data set to obtain estimates of the leverage for each metric.
5. Use survey measures such as Table 1 to obtain the risk discount factor for each metric.

6. Use Equation 13 to calculate the incremental change in each metric. Increase or decrease the emphasis on each metric as indicated.
7. Return to Step 3 periodically to update estimates of leverage and risk discount factors. Optimality is reached when the estimated incremental impact of changes in priorities approaches zero ($w_i^d \approx 0$), but periodic monitoring enables the system to adjust to changes in the environment.

Issues of Robustness

Adaptive control and linear models have proven robust in many social systems. We expect similar robustness here. For example, in Figure 2 we drew the tangent hyperplane as covering a small area relative to the total response curve. We did this because, formally, our methods are exact in an ϵ -neighborhood. However, in a real firm with real people our methods are at best an approximation. Fortunately, even if the variation about the operating point is large, the estimated directional change will still likely have the right sign. It will have the right sign even if some observations are on the opposite side of optimality relative to the operating point as long as the average is sufficiently far from optimal.

Furthermore, the adaptive control system will avoid false recommendations. In particular, the metrics thermostat only recommends changes when the regression estimate, $\hat{\lambda}_i$, is significantly different from zero. (The magnitude depends on RDF_i , but changes are recommended only when $\hat{\lambda}_i$ is non-zero.) If $\hat{\lambda}_i$ cannot be distinguished statistically from zero, then either (1) the metric is set optimally and no change is required (Figure 4), (2) the metric does not correlate with profit, or (3) the metric correlates with profit but there is not sufficient power in the estimation to measure this correlation. In the last two cases, the metrics thermostat does no harm – it does not recommend change. We recognize that the difference between Cases 1, 2, and 3 is important managerially. In Case 1, the metric is important and management has placed the right emphasis on the metric. It should not change the priority on the metric. In Case 2, the metric itself needs to be modified and, in Case 3, more data are needed. In our applications we have found that when $\hat{\lambda}_i$ cannot be distinguished from zero, management undertakes a detailed examination of the metric to distinguish Cases 1, 2, and 3.

The risk discount factor, RDF_i , is an interesting agency adjustment that has received much attention in the literature. In our case, RDF_i varies between 0 and 1, which means that it can reduce the emphasis on a metric by as much as a factor of $1/3 = 1/(1+2 \cdot 1) = 1/(1+2 \cdot RDF_i)$. Compared to the estimated leverage ($\hat{\lambda}_i$), which can vary by an order of magnitude, the adaptive control system is relatively robust to measurement errors or heterogeneity in RDF_i . Indeed, in our applications, the adjustment factor has varied between 40% and 80%. Nonetheless, we recommend that RDF_i be measured because (1) the measurement is relatively simple and (2) there is some, measurable impact.

Finally, we address an assumption made in our empirical applications. We examine what happens if Π_{ij}/M_{ij} varies dramatically? The result below gives us some idea of how the optimal weights vary if our homogeneity approximations are violated. The proof and the analytical equations are given in an appendix. Note that our homogeneity assumption applies to the ratio, Π_{ij}/M_{ij} . The assumption still holds if the separate variables, Π_{ij} and M_{ij} , vary as long as Π_{ij} and M_{ij} are correlated across projects. (Condition 4 below reproduces results in Baker [3] and Gibbons [18].)

RESULT. The increase in emphasis on a metric is larger if, for the same effort, (1) the expected increase in profit is large compared to the expected increase in the metric, (2) the expected signal-to-noise ratio is large, (3) the ability to increase the metric varies little across projects, and (4) increases in profit are correlated across projects with increases in the metric.

Application: Product Development Metrics and Profit

The metrics thermostat was developed and refined through an application with a multinational firm with almost \$20 billion in worldwide revenues. For the remainder of this article we call this firm, Tech. Working closely with senior technical management we selected a key division of Tech, which accounts for a significant fraction of Tech's revenues. This division sells complex office products in a business-to-business environment. Each of the twenty-two products launched in the last five years by this division share a core technology but differ in the details of electronics, materials, software, and mechanical design. This division earns revenue from the initial sales or rental of its products as well as sales of consumable supplies and service contracts.

Tech considers itself a sophisticated product development firm using state-of-the-art processes. Top management considers product development the competitive advantage of Tech and has invested heavily in tools, training, methods, culture, leadership, and incentives. Management is cognizant of the literatures about the drivers of successful product development. Tech's top management spent considerable effort trying to establish the right culture and incentives, including metrics. We adopt and study the metrics used by Tech and seek to provide recommendations to Tech on how to adjust their priorities with respect to these metrics. Naturally, other metrics might apply in different firms, or even different divisions of Tech.

We consider three top-level strategic metrics: customer satisfaction, time-to-market, and platform reuse. Customer satisfaction: Teams are strongly encouraged to use voice-of-the-customer methods, consider service during the design, identify the key vector of differentiation for their target market, use rigorous processes such as the house of quality, and take other actions to achieve customer satisfaction. Because these products are used in integrated office environments, compliance with office standards is likely to be a driver of customer satisfaction. Management believes that customer satisfaction will lead to long-term profit.

Time-to-market (TTM): Tech's US, Japanese, and European competitors have demonstrated success with rapid-deployment PD processes. To remain competitive, Tech invested heavily in an organized TTM process with clear goals, extensive training, incentives, a strong culture, and sufficient support. (The TTM process book alone is many inches thick.) The process encourages teams to do well on the TTM metric chosen by Tech. Team actions include the use of robust design methods, design for manufacturability, healthy relationships with suppliers, coordination among team members, and architecture designed for the easy integration of new technology. At the time of our application, Tech had just declared success in their TTM process and felt qualitatively that they were operating optimally. (We were told this after completing the measurement and estimation described in this section.)

Platform reuse: The nature of Tech's products is such that designs, and even some parts, can be reused. For example, software written for one product can be reused and modified for other products. In some cases, the user replaces products before parts wear out, especially in the rental market. With judicious design, Tech believed that some parts could be reused on new machines without any loss of customer satisfaction. At the time of this application, Tech was working hard to encourage platform reuse by adopting methods such as platform architecture design and reuse matrices.

These three strategic metrics best describe the implicit culture and reward system at this division of Tech. The teams were aware of the metrics and sought to do well with respect to the metrics by making the tradeoffs among actions that the teams perceived were in their best interests. Management believed the implicit reward system was in the best long-term interests of the firm. The culture in the division was strong, well understood, and, we believe, sufficiently homogeneous. Thus, Tech's current measures of the metrics described an operating point such as that in Figure 2b. (Tech's measures were based on extensive and significant managerial investments, including talking to many academics. They best represent the culture at Tech.) The goal of the metrics thermostat was to adjust the weights on these metrics, but a serendipitous outcome (beyond the scope of this article) was to focus attention on metrics. Future work might improve measurement of the metrics or identify new metrics.

In addition to the three strategic metrics, Tech's product development teams were measured on metrics that Tech felt would enable the team to succeed. These "enabling metrics" included the measurement (and implicit reward) of metrics such as standards compliance, rigor, market understanding, differentiation, and coordination. Top management at Tech believed that if the teams did well on these enabling metrics then they would do well on the three strategic metrics. Fortunately, the theory underlying the metrics thermostat applies equally well to estimate changes in the weights assigned to enabling metrics (as a means to achieve strategic metrics). In total, Tech used ten enabling metrics based on twenty-four specific measures. The specific enabling metrics were based on managerial judgment augmented with statistical analyses. We report on the system that Tech is now using and report Cronbach α 's for completeness and interpretation. Future research might improve the measures themselves.

Finally, to measure incremental profit, Equation 12 requires we include covariates that Tech believes affect profit, but which are outside the control of the PD teams. In cooperation with Tech's management, we selected seven covariates based on fourteen specific scales. Table 2 summarizes the strategic metrics, the enabling metrics, the covariates, and the underlying scales. Each of these measures has been previously identified by the academic literature as a driver of product development success.

Management was aware of the tradeoffs between the homogeneity requirements (tighter focus) and statistical power (broader focus). Based on their experience, they judged that the culture and incentives were sufficiently homogeneous within this division if we focused on all major new products launched in the last five years. Although this limits the number of data points with which to estimate Equation 12, management felt that expanding the data collection beyond the division might violate the necessary homogeneity assumptions.

With Tech's full cooperation, a product development graduate student sponsored by the National Science Foundation spent approximately nine months on site and/or in close contact with Tech. He was able to obtain detailed measures of profit (both short-term and expected long-term), strategic metrics, enabling metrics, and covariates. Team members completed a survey that provided Risk Discount Factors for each metric. The survey went through a number of pretests and iterations to the point where the PD team members felt they understood the RDF concept and felt that the survey answers reflected their risk preferences.

Among the sources of the data we were able to obtain were:

- Project proposals,
- Documents required by Tech's formal product development process,
- Business cases, phase review documents, and launch planning documents,
- Project post-mortems, technical review documents, and presentation slides,
- Competitive positioning diagrams,
- Detailed interviews with team members (with call-backs as necessary),
- Detailed interviews with management and with the finance group,
- Sales, service, and rental revenues,
- Manufacturing, delivery, and service costs,
- Judgments by senior-level managers about the expected long-term profit (1999 and beyond) for each product.

In total we collected reliable data on sixteen of the twenty-two PD projects. Because many of these data are actual "engineering" measures such as months, dollars, schedule time, etc., historic records were available for only these sixteen projects. Complete data could not be obtained for the other

six projects, either because documentation was not available or the team had disbanded and could not be reached. Tech now recognizes the need to collect and retain metrics in a common format. Managers have told us that this realization alone justifies the empirical project. Where appropriate, Table 2 reports Cronbach's α for the metrics categories. The α 's for the metrics average above 0.80. All α 's except those for "platform reuse" and "coordination" are 0.80 or higher. "Coordination" is above 0.70 and "platform reuse" is close to 0.70. The 0.80 meets Nunnally and Bernstein's [47, p. 265] criterion for comparing "groups" and is above their criterion (0.70) for exploratory research.⁷ For the purposes of this initial test application of the metrics thermostat, we feel these α 's are sufficient. Tech's management feels that this reliability testing is also a valued contribution. For example, they have initiated serious discussions about the scales underlying "platform reuse."

Table 2 also reports α 's for the covariates. Some of these α 's are lower than those for the metrics ("expected financial opportunity" and "team breadth and size"). Because neither of the low- α covariates made it into the final model, we tested the individual scales (after the model estimation). None would enter significantly. Nonetheless, these low α 's suggest future improvements in the basic measures by which Tech has chosen to manage.

Our data are broad in the number of variables, but limited to sixteen PD projects. Despite this limitation, we believe our data are unique in terms of a detailed look inside one division of an important multinational firm. Tech is considered a highly innovative firm with an excellent record in product development. Twenty-two major new-product projects in a five-year period is an impressive accomplishment. Our data are typical of what one can expect to obtain when the goal is to provide practical recommendations to a firm and illustrate why robustness was a concern in the development of the metrics thermostat.

To maintain confidentiality in this article, we rescaled the data such that the largest value of any metric, covariate, or profit measure is 5.0. Because Tech knows the scale of measurement, they need only multiply the weights (\hat{w}_i^d) we report by constants known to them. For the purpose of this article, we interpret the weights relative to the scaling disguise. It does not affect the qualitative interpretations.

Statistical Analysis

Regressions (Strategic Priorities and Covariates)

We begin with regressions that estimate leverage (the $\hat{\lambda}_i$'s). The sign of the recommended change in emphasis depends on $\hat{\lambda}_i$. (Its magnitude also depends on RDF_i .) A positive leverage suggests that Tech should increase its emphasis on that metric, while negative leverage implies that Tech should decrease its emphasis on that metric.

Recall from our discussion of robustness that changes are recommended only if $\hat{\lambda}_i$ is statistically different than zero – the thermostat does not recommend changes where none can be identified.. If $\hat{\lambda}_i$ is “zero” then extant methods are used to distinguish among the cases of optimality, no impact, and a need for greater precision. See Little [38] among others.

We examine two dependent variables: (1) immediate profit as measured by the profit per day that this product has earned for Tech since its launch,⁸ and (2) long-term profit as estimated by Tech’s management based on all the data they have collected to date about the product and its market. We focus on the strategic metrics (customer satisfaction, time-to-market, and platform reuse), but allow models that include direct effects from the enabling metrics. We were judicious in our choice of models, but tested the final models to determine whether any further variables (metrics or covariates) would enter with a significant t-statistic. They did not. In addition, we used interactive three-dimensional rotating plots to visually examine the data. These interactive plots helped protect our application from spurious interpretations. The regressions, the plots, and the implications were all discussed with Tech’s senior technical managers and judged to have high face validity.

Table 3 reports the results. Each of the three strategic metrics affects either short- or long-term profit. Customer satisfaction is strongly positive suggesting major profit potential from a shift toward more emphasis on customer satisfaction. On the other hand, both time-to-market and platform reuse are negative. The platform reuse coefficient is strongly negative suggesting that Tech has clearly overshot on this metric. The time-to-market coefficient is small compared to the coefficients for the other two strategic metrics suggesting that Tech has it about right with respect to time-to-market. Given all the approximations in the data and the analysis, Tech’s managers are not overly concerned with this small (but significant) negative coefficient.

In addition to the strategic metrics, one enabling metric (“consider service well”) has a direct effect on immediate profits. This is consistent with Tech management’s qualitative belief that they need to consider service more completely early in the product-design phase. In fact, a consortium of companies, of which Tech is a member, had already identified “service” as an important opportunity in the product development processes [52, p. 67]. The only covariate that enters these initial equations is “availability of resources.”

Leverage Regressions (Enabling Metrics)

Table 3 also reports regressions in which the strategic metrics serve as dependent variables and the enabling metrics (and covariates) are potential explanatory variables.⁹ These regressions reflect Tech’s belief that they can improve their strategic performance with processes that encourage excellence on the enabling metrics. The first regression suggests that Tech can improve its customers’ satisfaction if (1) the teams improve their compliance with regulatory, environmental, and industry stan-

dards, (2) the teams work with suppliers with which they are familiar, and (3) they consider service carefully in the design process. These interpretations were all judged by Tech's management to have high face validity.

The second regression (columns 8 and 9) suggests that Tech can improve its time-to-market if it works with familiar suppliers. It probably has the right strategic emphasis on other enabling metrics – not surprising given the recent corporate initiatives directed at getting to market faster.

The platform-reuse regression (columns 10 and 11) also makes intuitive sense.¹⁰ Working with suppliers in which Tech has confidence and avoiding differentiation will improve platform reuse. Considering service leads to greater platform reuse – probably because proven parts are easier to service. The negative coefficient on “rigor” is interesting. Because platform reuse has a negative effect on profit, it appears that more rigorous product development tools indicate less platform reuse. This decreased platform reuse leads to greater profits.

Changes in the Weights

Table 4 (second column) reports the average Risk Discount Factor (RDF_i) for each metric. Because some metrics have both direct and indirect effects we use all the regressions in Table 3 to calculate the net weight of each metric. For example, “consider service well” has a direct effect on short-term profit as well as an indirect effect through platform reuse. It has an indirect effect on long-term profit through customer satisfaction. To retain confidentiality and avoid revealing the exact scaling of short-term and long-term profit, we simply sum the 5-point measures to obtain “total” profit. This is sufficient for qualitative interpretations. The end result of this arithmetic gives us an estimate of the net leverage. We then divide that estimate by $(1 + 2RDF_i)$ to obtain the strategic priorities (w_i^d 's) implied by Equation 13. These are given in the third column of Table 4. Based on these weights, customer satisfaction is a strategic opportunity. Perhaps Tech's recent initiatives on time-to-market and platform reuse have caused its teams to lose sight of customer satisfaction. This is consistent with our own qualitative experience at Tech. (See related discussions in Deck [14].) On the other hand, Tech should seriously reconsider their latest platform-reuse initiatives. Investing \$100 million in platform reuse may have negative profit implications. As one consultant suggested to us, platform reuse at Tech might have gone so far as to become the opposite of innovation.

At a more-detailed level, Table 4 suggests that Tech should seek better vectors of differentiation, put more emphasis on the standards compliance, and encourage the greater use of formal product development tools. The recommendations with respect to suppliers are mixed. On one hand, there are advantages to working with familiar suppliers. On the other hand this might be encouraging too much platform reuse (and thus inhibiting innovation and lowering profits).

Tech's Reaction

After reviewing the results in Table 4, Tech's management told us that the results were consistent with suspicions they had had about their product development process. The metrics thermostat had quantified their intuition. The TTM process was considered a major success, but they had hoped that TTM did not sacrifice customer satisfaction. Platform reuse was important, but might be easier for the PD teams to achieve than customer satisfaction. The teams might have overshot on this metric. Because these results were consistent with other inputs, they acted.

To enhance customer satisfaction, Tech created the role of a "marketing engineer." A marketing engineer is a full member of each new PD team. He or she is responsible for assuring that the voice of the customer is incorporated in the design and that the product is designed to be marketed. Tech has also adjusted its channels to reach the customer better and, in particular, to match customers with the appropriate products. They are working with their technical representatives in the channel to enhance service (a key enabling metric in Table 4).

In addition, Tech is undertaking a major study of platform reuse to optimize their portfolio with respect to upstream/downstream technological development, a balance of product variants and major redesigns, and enterprise coherence in software development (standards compliance in Table 4).

Consistent with the philosophy of dispersed product development, Tech's management is trying to change the culture and incentives that determine the implicit weights on the metrics. They are also investing in new PD tools and in educating the teams. However, as they change the culture and implicit incentives, they are leaving the detailed decisions on actions to achieve the metrics to the self-managed, empowered teams.

Toward Lean Metrics

Tech follows a complex and thorough product development process. One of the complaints that we have heard from team members at Tech is that the PD process is too complete – they are not sure where they should place emphasis. While strategic metrics set and communicate overall priorities, the value of the enabling metrics must be weighed carefully against the cost of collecting data on these metrics. One side benefit of our analysis is that we have helped to identify a set of useful enabling metrics. Tech believes that the overall strategic measures of customer satisfaction, time-to-market, and platform reuse, with perhaps the addition of service consideration, appear to be a simple and relatively complete set of strategic metrics. At minimum, they are good correlates of incremental profits at the operating point. If these metrics continue to be useful in further updates with adaptive control, then they might begin to form the basis of a set of "lean" metrics.

Summary and Future Directions

In the last 10-15 years many large firms have invested heavily in strategic initiatives – in many cases spending over \$100 million on each initiative. They have implemented and encouraged these initiatives with metrics. Besides customer satisfaction, time-to-market, and platform-reuse, firms have invested in just-in-time inventory management, Kaizen methods, core-competence identification, re-engineering, strategic portfolio selection, and other techniques that have been heralded to improve competitiveness. Most initiatives lead to early improvements but later disappoint. This should not surprise us. Too much of a good thing can be bad and initiatives once begun are hard to stop. For example, Boulding, Morgan and Staelin [6], Simonson and Staw [55], and Staw [62] suggest that it is difficult to de-escalate actions once they've begun.

In this article we illustrated a practical adaptive control method to adjust priorities on a firm's chosen metrics. The system was designed to address the issues important to management and to work within the data limitations imposed by a focus on one or more divisions of a single firm. The system was designed to be robust with respect to these data limitations and to the tendency of PD teams to act in their own best interests. We have striven to make tradeoffs between rigor and practicality. The initial analyses at Tech show promise and we have begun applications at a major automobile manufacturer. We are also experimenting with modifications of the methodology to address procurement (U.S. Navy) and sustainment (USAF).

Naturally, our methods can be improved. At Tech we made tradeoffs with respect to focus (homogeneity within a division) versus breadth (multiple heterogeneous divisions). While we were careful in our initial statistical analyses, we will gain confidence and understanding through further applications which will lead to continuous improvement of the metrics thermostat and will suggest generalized implications. There might be practical ways to improve the approximations within the hyperplane, improve the measurement of RDF_i , and combine experimentation with statistical estimation.

We recognize that adjusting the emphasis on metrics is only one part of the overall system as described in Figure 1. Our focus fits within the overall process of selecting the right metrics, establishing the culture, providing new methods and processes to the teams, enhancing communication among team members (and between teams), studying informal and formal incentive systems, and a variety of other important research areas in product development. We hope our contribution to fine-tuning metrics improves that overall process.

References

1. Atuahene-Gima, Kwaku (1995), "An Exploratory Analysis of the Impact of Market Orientation on New Product Performance: A Contingency Approach," *Journal of Product Innovation Management*, 12, 275-293.
2. Ayers, Doug, Robert Dahlstrom, and Steven J. Skinner, "An Exploratory Investigation of Organizational Antecedents to New Product Success," *Journal of Marketing Research*, 34, 1, (February), 107-116.
3. Baker, George (1992), "Incentive Contracts and Performance Measurement," *Journal of Political Economy*, 100, 598-614.
4. Bergen, Mark, Shantanu Dutta, and Orville C. Walker, Jr. (1992), "Agency Relationships in Marketing: A Review of the Implications and Applications of Agency and Related Theories," *Journal of Marketing*, 56, (July), 1-24.
5. Bonner, Joseph, Robert Ruckert and Orville Walker (1998), "Management Control of Product Development Projects," Marketing Science Institute Working Paper 98-120.
6. Boulding, William, Ruskin Morgan, and Richard Staelin (1997), "Pulling the Plug to Stop the New Product Drain," *Journal of Marketing Research*, 34, (February), 164-176.
7. Calantone, Roger J. and C. Anthony di Benedetto (1988), "An Integrative Model of the New Product Development Process: An Empirical Validation," *Journal of Product Innovation Management*, 5, 201-215.
8. Chan, Christine W. Y. (1999), "Measuring Non-Monetary Incentives Using Conjoint Analysis," Electrical Engineering and Computer Science Master's Thesis, Massachusetts Institute of Technology, Cambridge, MA 02139 (June).
9. Cooper, Robert G. and Ulricke de Brentani (1991), "New Industrial Financial Services: What Distinguishes the Winners," *Journal of Product Innovation Management*, 8, 75-90.
10. _____ and Elko J. Kleinschmidt (1993), "Major New Products: What Distinguishes the Winners in the Chemical Industry?," *Journal of Product Innovation Management*, 10, 90-111.
11. _____ and _____ (1994), "Determinants of Timeliness in Product Development," *Journal of Product Innovation Management*, 11, 381-396.
12. _____ and _____ (1995), "Benchmarking the Firm's Critical Success Factors in New Product Development," *Journal of Product Innovation Management*, 12, 374-391.
13. Datar, Srikant, C. Clark Jordan, Sunder Kekre, Surendra Rajiv, and Kannan Srinivasan (1997), "Advantages of Time-Based New Product Development in a Fast-Cycle Industry," *Journal of Marketing Research*, 34, 1, (February), 36-49.
14. Deck, Mark (1999), "Product Metrics," *Visions*, 23, 1, (January), 20 and 25.
15. Eliashberg, Jehoshua and John R. Hauser (1985), "A Measurement Error Approach for Modeling Consumer Risk Preference," *Management Science*, Vol. 31, No. 1, (January), 1-25.
16. Eppinger, Steven D. (1998), "Information-Based Product Development," Presentation to the Research Advisory Committee of the Center for Innovation in Product Development, MIT, Cambridge, MA (January 12).
17. _____, Daniel E. Whitney, Robert P. Smith, and David A. Gebala (1994), "A Model-Based Method for Organizing Tasks in Product Development", *Research in Engineering Design*. 6, 1, 1-13.

18. Gibbons, Robert (1997), "Incentives and Careers in Organizations," in D. Kreps and K. Wallis (Eds.) *Advances in Economic Theory and Econometrics*, Cambridge, UK: Cambridge University Press, 1-37.
19. Goldenberg, Jacob, Donald R. Lehmann, and David Mazursky (1999), "The Primacy of the Idea Itself as a Predictor of New Product Success," Marketing Science Institute Working Paper, 99-110.
20. Gonzalez-Zugasti, Javier, Kevin. Otto, and J. Baker (1998), "A Method for Architecting Product Platforms with an Application to Interplanetary Mission Design," *1998 ASME Design Automation Conference*, Atlanta, GA.
21. Griffin, Abbie (1997), "The Effect of Project and Process Characteristics on Product Development Cycle Time," *Journal of Marketing Research*, 34, 1, (February), 24-35.
22. _____ (2000), "Product Development Cycle Time for Business-to-Business Products," Working Paper, (Urbana-Champaign, IL: University of Illinois), October.
23. _____ and Albert L. Page (1993), "An Interim Report on Measuring Product Development Success and Failure," *Journal of Product Innovation Management*, 10, 291-308.
24. _____ and _____ (1996), "PDMA Success Measurement Project: Recommended Measures for Product Development Success and Failure," *Journal of Product Innovation Management*, 13, 478-496.
25. _____ and John R. Hauser (1993), "The Voice of the Customer," *Marketing Science*, 12, No. 1, (Winter), 1-27.
26. _____ and _____ (1996), "Integrating Mechanisms for Marketing and R&D," *Journal of Product Innovation Management*, 13, 3, (May), 191-215.
27. Hauser, John R. and Don Clausing (1988), "The House of Quality," *Harvard Business Review*, Vol. No. 3, (May-June), 63-73.
28. _____ and Gerry Katz (1998), "Metrics: You are What You Measure!" *European Management Journal*, 16, 5, 516-528.
29. Holmes, Maurice (1999), "Product Development in the New Millennium - a CIPD Vision," Product Development Management Association Conference, Marco Island (October).
30. Holmstrom, Bengt (1979), "Moral Hazard and Observability," *Bell Journal of Economics*, 10, 74-91.
31. _____ and Paul Milgrom (1987), "Aggregation and Linearity in the Provision of Intertemporal Incentives", *Econometrica* 55, (March), 303-28.
32. Ittner, Christopher D. and David F. Larcher (1997), "Product Development Cycle Time and Organizational Performance," *Journal of Marketing Research*, 34, 1, (February), 13-23.
33. Kahn, Kenneth B. (1996), "Interdepartmental Integration: A Definition with Implications for Product Development Performance," *Journal of Product Innovation Management*, 13, 137-151.
34. Keeney, Ralph L. and Howard Raiffa (1976), *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*, (New York: John Wiley & Sons).
35. LaFountain, Burt D. (1999), "An Empirical Exploration of Metrics for Product Development Teams," "Mechanical Engineering Master of Science Thesis, Massachusetts Institute of Technology, Cambridge, MA 02139 (February).
36. Lambert, Denis and Stanley F. Slater (1999), "Perspective: First, Fast, and On Time: The Path to Success. Or is it?," *Journal of Product Innovation Management*, 16, 427-438.

37. Lee, Mushin and Dohyeong Na (1994), "Determinants of Technical Success in Product Development When Innovative Radicalness is Considered," *Journal of Product Innovation Management*, 11, 62-68.
38. Little, John D.C. (1966), "A Model of Adaptive Control of Promotional Spending," *Operations Research*, 14, 6, (November-December), 1075-1098.
39. Luenberger, David G. (1969), *Optimization by Vector Space Methods*, (New York, NY: John Wiley & Sons, Inc.)
40. Menezes, Melvyn A. J. (1994), "Xerox Corporation: The Customer Satisfaction Program (A)," *Harvard Business School Case 9-594-109*, (Boston, MA: Publishing Division, Harvard Business School).
41. Mirrlees, James (1974), "Notes on Welfare Economics Information and Uncertainty," in Balch, McFadden and Wu (Eds.), *Essays on Economic Behaviour Under Uncertainty*, (Amsterdam: North-Holland), 243-257.
42. Mishra, Sanjay, Dongwook Kim, and Dae Hoon Lee (1996), "Factors Affecting New Product Success: Cross-Country Comparisons," *Journal of Product Innovation Management*, 13, 530-550.
43. Montoya-Weiss, Mitzi and Roger Calantone (1994), "Determinants of New Product Performance: A Review and Meta-Analysis," *Journal of Product Innovation Management*, 11, 397-417.
44. Moore, William L., Jordan J. Louviere and Rohit Verma (1999), "Using Conjoint Analysis to Help Design Product Platforms," *Journal of Product Innovation Management*, 16, 1, (January), 27-39.
45. Moorman, Christine (1995), "Organizational Market Information Processes: Cultural Antecedents and New Product Outcomes," *Journal of Marketing Research*, 32, (August), 318-335.
46. _____ and Anne S. Miner (1997), "The Impact of Organizational Memory on New Product Performance and Creativity," *Journal of Marketing Research*, 34, 1, (February), 91-106.
47. Nunnally, Jum C. and Ira H. Bernstein (1994), *Psychometric Theory*, (New York, NY: McGraw-Hill Inc.).
48. Nobeoka, Kentaro and Michael A. Cusumano (1997), "Multiproject Strategy and Sales Growth: The Benefits of Rapid Design Transfer in New Product Development," *Strategic Management Journal*, 18, 3, 169-186.
49. Olson, Eric M., Orville C. Walker, Jr., and Robert Ruckert (1995), "Organizing for Effective New Product Development: The Moderating Role of Product Innovativeness," *Journal of Marketing*, 59, (January), 48-62.
50. Phillips, Stephen, Amy Dunkin, James Treece, and Keith H. Hammonds (1990), "King Customer: At Companies That Listen Hard and Respond Fast, Bottom Lines Thrive," *Business Week*, (March 12), 88-94.
51. Rosen, Barry Nathan, Steven P. Schnaars and David Shani (1988), "A Comparison of Approaches for Setting Standards for Technological Products," *Journal of Product Innovation Management*, 5, 2, (June), 129-139.
52. Seering, Warren (1998), "Annual Report – Second Year," Center for Innovation in Product Development, M.I.T., Cambridge, MA (October).
53. Sharda, Ramesh, Gary L. Frankwick, Atul Deosthali, Ron Delahoussaye (1999), "Information Support for New Product Development Teams," Marketing Science Institute Working Paper, 99-108.
54. Simester, Duncan I., John R. Hauser, Birger Wernerfelt, and Roland T. Rust (2000), "Implementing Quality Improvement Programs Designed to Enhance Customer Satisfaction: Quasi-experiments in the U.S. and Spain," *Journal of Marketing Research* (February).

55. Simonson, Itamar and Barry M. Staw (1992), "De-escalation Strategies: A Comparison of Techniques for Reducing Commitment to Losing Courses of Action," *Journal of Applied Psychology*, 77, 4, 419-426.
56. Smith, Preston G. and Reinertsen, Donald G. (1991), *Developing Products in Half the Time*, (New York: Van Nostrand Reinhold).
57. _____ and _____ (1998), *Developing Products in Half the Time: New Rules, New Tools*, (New York: John Wiley & Sons)
58. Song, X. Michael, C. Anthony di Benedetto, and Yuzhen Lisa Zhao (1999), "Does Pioneering Advantage Exist?: A Cross-national Comparative Study," Marketing Science Institute Working Paper, 99-111.
59. _____, Mitzi M. Montoya-Weiss, and Jeffrey B. Schmidt (1997), "Antecedents and Consequences of Cross-functional Cooperation: A Comparison of R&D, Manufacturing, and Marketing Perspectives," *Journal of Product Innovation Management*, 14, 35-47.
60. _____ and Mark E. Parry (1997a), "The Determinants of Japanese New Product Successes," *Journal of Marketing Research*, 34, 1, (February), 64-76.
61. _____ and _____ (1997b), "A Cross-National Comparative Study of New Product Development Processes: Japan and the United States," *Journal of Marketing*, 61, (April), 1-18.
62. Staw, Barry M. (1976), "Knee-Deep in the Big Muddy: A Study of Escalating Commitment to a Chosen Course of Action," *Organizational Behavior and Human Performance*, 16, 27-44.
63. Suro, Roberto (2000), "For U.S. Aviators, Readiness Woes are a 2-Front Struggle; Veteran Mechanics and Spare Parts in Short Supply," *Washington Post*, (February 3), A04.
64. Ulrich, Karl T. and Steven D. Eppinger (1995), *Product Design and Development*, (New York, NY: McGraw Hill).
65. Witter, Jerome, Don Clausing, Ludger Laufenberg and Ronaldo Soares de Andrade (1994), "Reusability – The Key to Corporate Agility," *Proceedings of the Conference on New Approaches to Development and Engineering*, Gothenburg, Sweden, (May).

Appendix: Proof of Comparative Statics Result

Using the notation of the text, we obtain:

$$(A1) \quad \textit{division profits} = E_j \left[\sum_{j=1}^J \sum_{i=1}^n \left\{ \Pi_{ij} e_{ij} - \frac{1}{2} C_{ij}^2 (e_{ij})^2 - \frac{1}{2} r \sigma_i^2 (w_i^d)^2 + \textit{error}_j^\pi \right\} \right]$$

For a given set of division priorities, w_i^d 's, each team will select its efforts to maximize its certainty equivalent. Recognizing that publicly traded firms should be risk neutral, we substitute these optimal efforts into Equation A1 and cancel terms to obtain:

$$(A2) \quad \textit{division profits} = \sum_{j=1}^J \sum_{i=1}^n \left\{ \frac{w_i^d M_{ij} \Pi_{ij}}{C_{ij}^2} - \frac{(w_i^d)^2 M_{ij}^2}{2C_{ij}^2} - \frac{1}{2} r \sigma_i^2 (w_i^d)^2 \right\}$$

Differentiating Equation A2 with respect to w_i^d and recognizing that $(1/J)$ times the sum over projects is just the (empirically estimated) expected value, we obtain:

$$(A3) \quad w_i^d = \frac{E_j[\Pi_{ij} M_{ij} / C_{ij}^2]}{E_j[M_{ij}^2 / C_{ij}^2] + r \sigma_i^2}$$

Finally, if the unobserved cost structure (C_{ij}^2) is homogeneous, we use the definitions of variance and covariance to derive an expression for the strategic priorities as a function of definable, but hard-to-measure, quantities:

$$(A4) \quad w_i^d = \frac{\frac{E_j[\Pi_{ij}]}{E_j[M_{ij}]} + \frac{\text{cov}(M_{ij}, \Pi_{ij})}{E_j^2[M_{ij}]}}{1 + \frac{\text{var}(M_{ij})}{E_j^2[M_{ij}]} + \frac{r \sigma_i^2 C_{ij}^2}{E_j^2[M_{ij}]}}$$

Inspection of Equation A4 establishes Result 1

Table 1. EXAMPLE SURVEY FOR THE RISK DISCOUNT FACTOR (Technically, 1 - RDF)

Imagine that you are on a team about to embark on a project to design and develop a new product. This will be a balanced cross-functional team, but you do not yet know who the other members will be.

You will receive many rewards--above and beyond your salary--based on your team's performance. Some of these might include:

- Monetary bonuses
- Promotion
- Respect from colleagues
- Opportunities to work on interesting projects in the future

Some of the above mentioned rewards may be explicit—for example, they are formally determined by contracts with management. Others are implicit—the reward structure exists within the culture of your firm. In this survey, we ask you to consider ALL the rewards, explicit and implicit, that you might receive based on your team's performance on a product development project.

On the following page are several aspects of your team's performance that might be judged, observed, or measured by others to determine the rewards (explicit *and* implicit) that you receive. You and your team have the opportunity to impact these aspects of performance through your efforts.

For each aspect of performance on the list, imagine that you determine *up front* how much effort you will expend to affect it. Do not worry about determining what that effort would actually be. Then consider the following two scenarios. They differ in how your explicit and implicit rewards are determined. *Note that for each aspect of performance you would choose the same amount of effort for Scenario B as for Scenario A.*

Scenario A: *You decide how much effort to put in to the aspect of performance. Your reward is based on the judgment or measurement of the aspect of performance by someone outside the team. You cannot be certain what the judged or measured value of your performance will actually be, therefore the amount of reward you will receive is not certain. For the amount of effort that you have chosen to allocate, there is some average expected reward that could be calculated across many projects of the same type. However, there is uncertainty for any individual (i.e., your) project.*

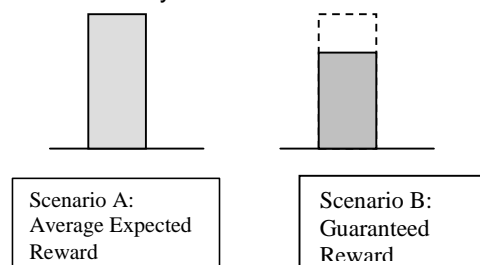
Scenario B: *You allocate the same amount of effort as in Scenario A to the aspect of performance. However, the amount of reward you receive for this is determined in advance. There is no uncertainty.*

If the guaranteed rewards from Scenario B were equal to or greater than the average expected rewards from Scenario A, most people would prefer Scenario B because Scenario B eliminates risk. In fact, some people would prefer Scenario B even if the guaranteed rewards were less than the average expected rewards from Scenario A.

For each aspect of performance, we would like you to answer the following:

At what value of the guaranteed rewards from Scenario B (as a percentage of the average respected reward from Scenario A) would you be indifferent between the two scenarios?

For example, you might prefer Scenario B if the guaranteed rewards were equal to 99 percent of the average expected rewards from Scenario A. But you might prefer Scenario A if the guaranteed rewards were equal to 1 percent of the average expected rewards from Scenario A. Thus there would be some percentage between 1 and 99 for which you would be indifferent between the two scenarios.



**Table 2. Detailed Measures of Metrics and Covariates
(numbers in parentheses are Cronbach α 's)**

Strategic Metrics		Enabling metrics (continued)	
Time to market	Slip rate (% of schedule overshoot)	Differentiation (.90)	Degree of differentiation from competitive products
Customer satisfaction	Survey measure of post-install satisfaction		Degree of differentiation from Tech's products
Platform reuse (.65)	Product leverages platform elements well	Coordination (.71)	Level of coordination achieved within team
	Reuse of previous product elements		Number of critical issues assessed at phase review
	Expected reuse from this product		Level of coordination achieved between team and internal value chain partners
			Number of major issues assessed at phase review
			Quality of integrated plan
Enabling Metrics		Covariates	
Feasible vector of differentiation	Product achieves intended vector of differentiation	Availability of resources (.79)	Resources available for continuance
Standards compliance (.80)	Attention to regulatory, environmental standards		Skills available for continuance
	Compliance with industry standards		Fits Tech's corporate image
Rigor (.86)	Use Quality Function Deployment	Fits corporate image	Aligns w/ corporate strategy and core competence
	Use robust design practices	Core competence alignment (.84)	Product is grounded in the market attack plan
	Design for manufacturing capability		Size of strategic market advantage to be gained
Supplier confidence and health (.91)	Health of relationships with existing suppliers	Size of strategic opportunity (.84)	Size of strategic technology advantage to be gained
	Confidence in delivery by suppliers already selected		Product acquisition spending
Supplier maturity (.88)	Degree to which suppliers have been selected early	Expected financial opportunity (.55)	Expected product placement
	Maturity of relationship with suppliers		Expected revenue
Consider service well	Design process considers service well		Expected lifetime profit
Understand needs and market (.86)	Effort to gather the voice of the customer		
	Effort undertaken to study market characteristics		
Technological advantage (.84)	Technology advantage with this product	Team size and breadth (.61)	Core team size
	Richness of technology options		Core team dispersion
	Architecture allows easy integration of technology	Experience	Tech's experience in this market
	Newness of technology		

Table 3. Leverage Regressions ($\hat{\lambda}_i$)

	Profit to date		Long-term profit		Customer satisfaction		Time to market		Platform re-use	
	<i>Coeff.</i>	<i>t-stat.</i>	<i>Coeff.</i>	<i>t-stat.</i>	<i>Coeff.</i>	<i>t-stat.</i>	<i>Coeff.</i>	<i>t-stat.</i>	<i>Coeff.</i>	<i>t-stat.</i>
Strategic metrics										
Customer satisfaction			4.14	5.0						
Time to market			-0.48	-2.2						
Platform reuse	-2.65	-6.9								
Enabling metrics										
Vector of differentiation									-0.46	-2.1
Rigor (QFD, robust design)									-0.31	-3.5
Standards compliance					0.41	3.3				
Supplier confidence, health									0.63	3.3
Supplier maturity					0.19	4.5	0.75	2.6		
Consider service well	0.78	2.7			0.22	4.7			0.25	2.3
Covariates										
Availability of resources	1.36	3.9								
Fits corporate image									-0.41	-4.7
R^2 (<i>F-statistic</i>)	0.85	20.2	0.66	12.5	0.88	28.4	0.33	6.8	0.83	9.4
Adjusted R^2	0.81		0.61		0.85		0.28		0.74	

Table 4. RDF and the Implied Changes in Emphasis on Metrics (w_i^d)

	RDF <i>(survey measure)</i>	Implied change in emphasis to im- prove profits
Profits		
Profit to date		
Long-term profit		
Strategic metrics		
Customer satisfaction	0.27	2.7
Time to market	0.25	-0.3
Platform reuse	0.34	-1.6
Enabling metrics		
Vector of differentiation	0.28	0.8
Rigor (QFD, robust design)	0.22	1.1
Standards compliance	0.22	0.6
Supplier confidence, health	0.38	-1.0
Supplier maturity	0.38	0.2
Consider service well	0.22	0.2
Understand needs, market	0.36	0.0

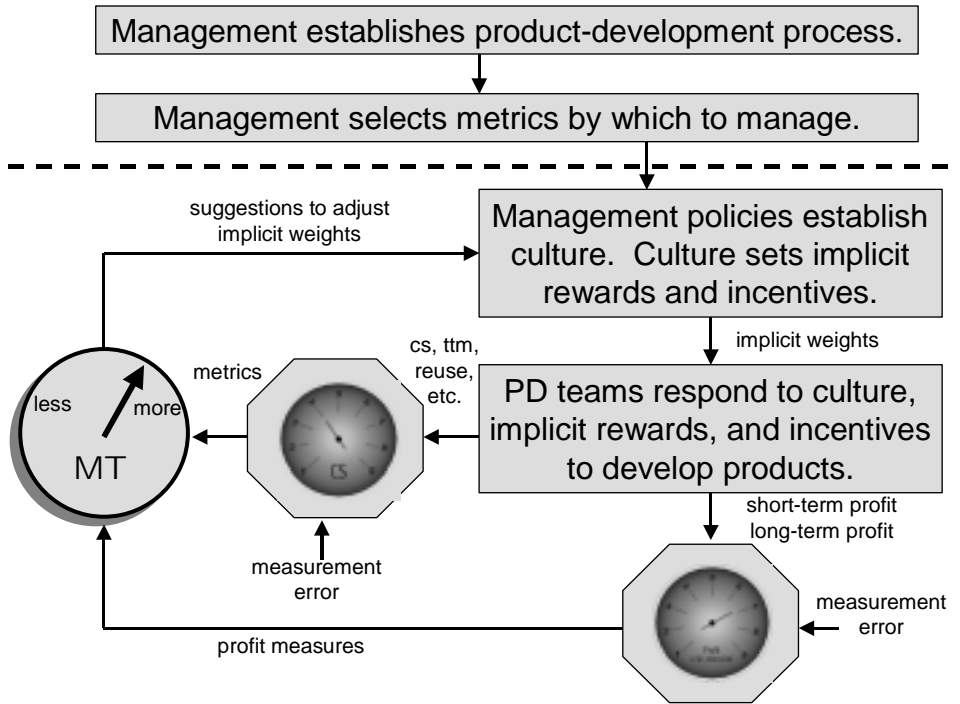


Figure 1. Managing a Dispersed Product Development Process with Metrics
 (The focus of the metrics thermostat, MT, is below the dotted line.)

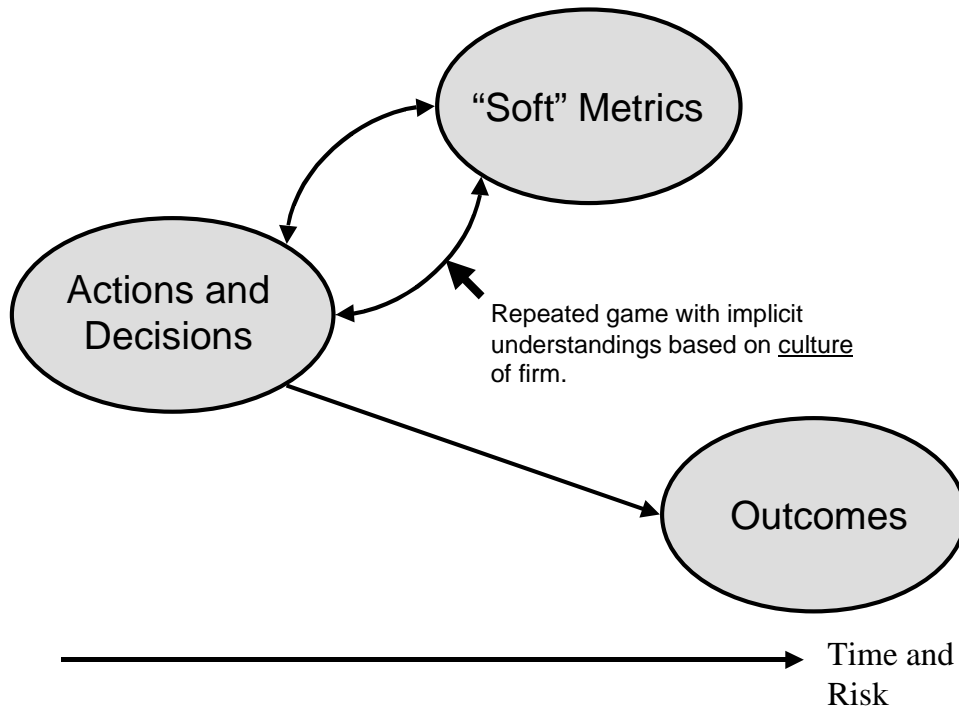


Figure 3. Performance Measure Interpretation of Metrics

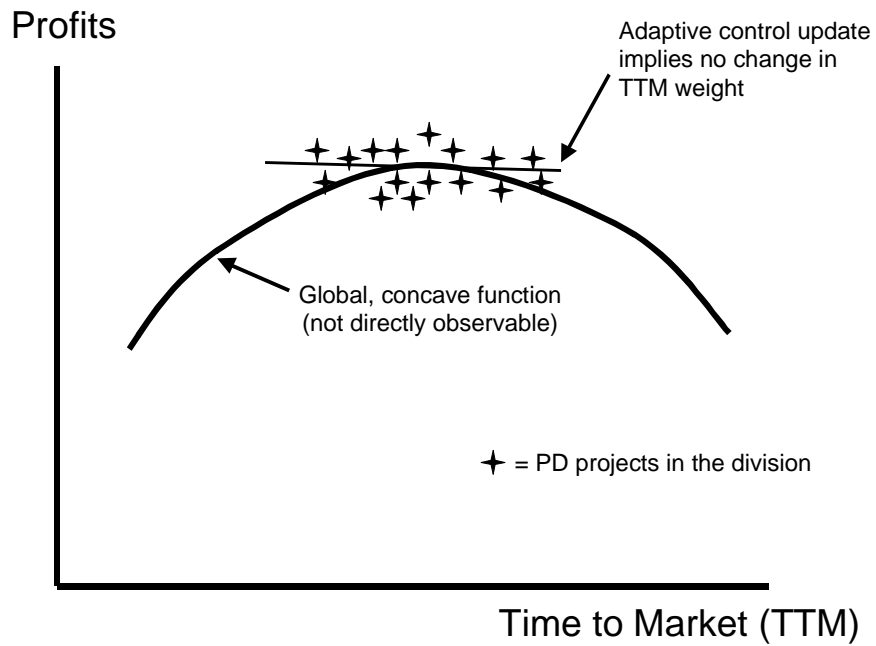
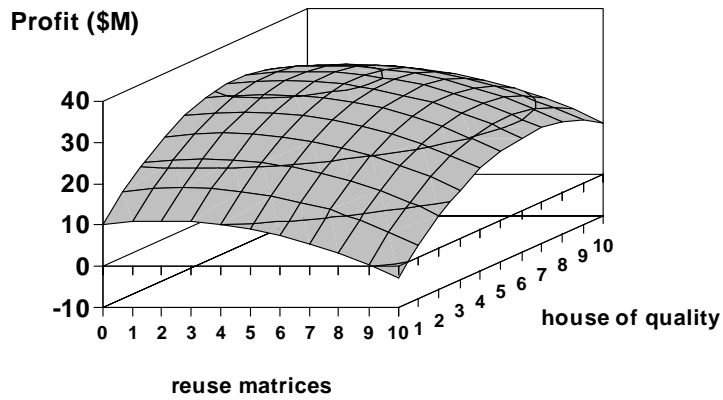
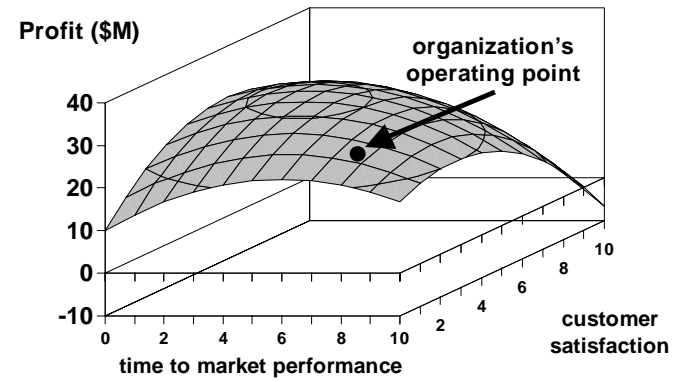


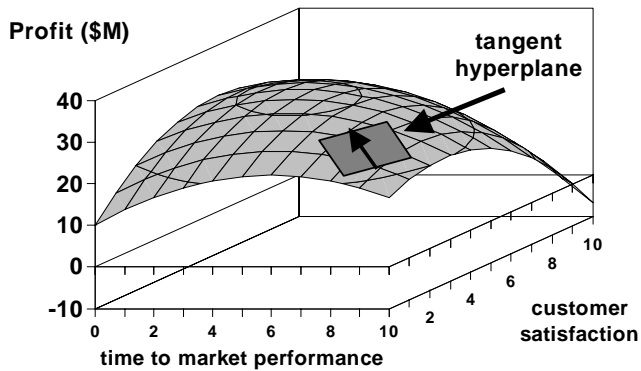
Figure 4. If the leverage of TTM is zero, then one explanation is that the emphasis on TTM is optimal and should not be changed.



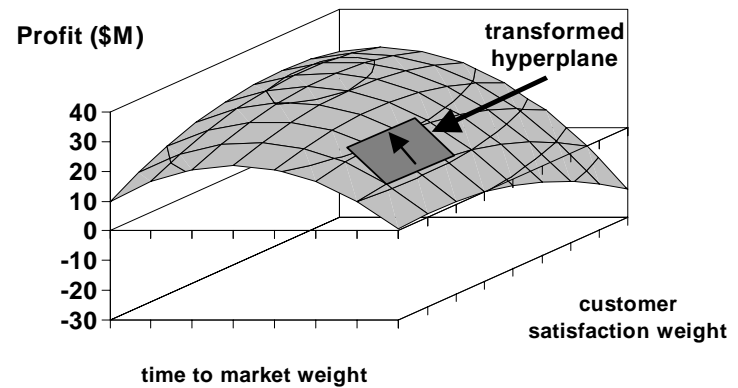
a) Team takes actions. Shown are two of the many actions.



b) Team's actions lead to values of the metrics.



c) Incremental improvements based on hyperplane approximation.



d) Changes in metrics are transformed to changes in the weights on metrics.

Figure 2. Illustrative Example of Adaptive Control

Endnotes

¹ This is a subtle point. While the mapping from actions to metrics is many-to-one, the team will choose to minimize their costs for given levels of metrics. This optimization by the teams produces an envelope in action space that can be one-to-one. Even if the mapping were not one-to-one, the firm would be indifferent between alternative actions that produce the same metrics at the same cost to the team and the firm.

² The plot in Figure 2a is generated with $\pi = 10 + (132a_1 + 36a_2 - 5a_1^2 - 9a_2^2)/42$. Equation 1 transforms the plot to $\pi = 10 + (36m_1 + 24m_2 - 3m_1^2 - 2m_2^2 - 2m_1m_2)/6$.

³ A few technical properties are sufficient to guarantee optimality. For example, optimality is guaranteed if the step size is chosen such that the mapping is a contraction mapping (Luenberger [39], p. 272). Empirically, these properties are likely to be satisfied. Even if the surface is multi-modal, this procedure can work if we begin with a large step size and then move to a small step size. For those readers who enjoy the mathematics, we suggest trying the concept with the equations in footnote 2 starting with $m_1 = 8$ and $m_2 = 0$. By taking derivatives analytically and taking small steps along the gradient vector one can reach 99% of optimal in two steps. This increases profits by 21%. With a more extreme starting point, $m_1 = 0$ and $m_2 = 10$, one can also achieve the 99% goal in two steps while increasing profit by 88%.

⁴ The certainty equivalent is the level of guaranteed rewards that the team would accept as equivalent to the risky rewards in Equation 2. See [34, p.161] for more details. Constant risk aversion is a reasonable approximation in the tangent hyperplane. When the utility function is constantly risk averse, the risk terms do not depend on the operating point. The team incurs the risk of measurement only for incremental changes in the weights and metrics. Thus, we need no risk term that depends on the square of the sum of w_i^o and w_i . If the metrics are measured far in the future, the agent might discount them. We can model this by adding a discount factor, $\gamma_i < 1$, for each metric.

⁵ For example, if we were to include $\partial c^o / \partial e_{ij}^o$, then it would drop out in the derivation of Equation 8 and, later, Equation A3. We also assume that the cross-partial derivatives of c^o are small compared to C_{ij}^2 . This is an approximation, but one that we feel is justified by the empirical problem.

⁶ Equation 12 differs from the linear regressions common in the antecedents and consequences literature. In that literature, the goal is to identify those metrics which vary across firms and influence profit. Equation 12 is used within a firm to identify incremental improvements in weights rather than global optima.

⁷ Nunnally and Bernstein [47] apply when comparing groups of people (0.80 criterion) or exploring scales to measure psychological constructs (0.70 criterion). For these purposes pushing α beyond 0.80 is “wasteful of time and money.” However, unlike laboratory research in psychology, our equations (RDF) explicitly adjust for uncertainty in the metrics. Nonetheless, we must be careful when interpreting reliability measures for noisy metrics.

⁸ We use profit per day (total profit/days since launch) to adjust for the fact that the sixteen products are the result of a continuous product development process spread over five years. Some have been on the market since 1994; some have been launched recently.

⁹ The metrics “technology advantage,” “differentiation,” and “coordination” do not enter the models nor do the covariates “core competence alignment,” “size of strategic opportunity,” “expected financial opportunity,” “team size and breadth,” and “experience.”

¹⁰ “Technology advantage” is collinear with “vector of differentiation,” “rigor,” “supplier confidence and health,” and “consider service well.” The regression of “technology advantage” on these variates has an R^2 of 0.69. With more data or less collinear data, “technology advantage” might enter the equation.