CAMBRIDGE

Cambridge University Press
0521812437 - Rethinking Performance Measurement: Beyond the Balanced Scorecard
Marshall W. Meyer
Excerpt
More information

Introduction

DISSATISFACTION with performance measurement systems runs high. Many firms, perhaps the majority, suspect that they haven't got it right. A 1995 article in *Chief Financial Officer* begins, "According to a recent survey, 80 percent of large American companies want to change their performance measurement systems . . ."[1] Unsurprisingly, the turmoil in performance measurement is ongoing. Startup companies struggling for capital must continually adjust their metrics.[2] And it is commonplace for large firms to undertake annual overhauls of their performance measurement systems.[3]

Why the turmoil and dissatisfaction? One cause is the ongoing search for non-financial predictors of financial performance: "Yesterday's accounting results say nothing about the factors that actually help grow market share and profits – things like customer service innovation, R&D effectiveness, the percent of first-time quality, and employee development."[4] Another cause, ironically, is a surfeit of measures: many corporate controllers cite the burdens imposed by "newfangled performance measures" as a key source of burnout.[5] Anecdotal reports such as these suggest that executives are seeking measures that controllers and chief financial officers have so far been reluctant or unable to deliver. The result is frustration on both sides.

Whether the problem is too few or too many measures, many accountants believe that corporate performance measurement systems do not support management objectives well. According to the Institute of Management Accountants, the proportion of accountants rating their performance measures as "poor" or "less than adequate," the bottom two categories on a six-point scale where the fourth category is "adequate," has remained substantial, ranging from 35 percent in 1992 to 43 percent in 1993, 38 percent in 1995, 43 percent in 1996, 34 percent in 1997, 40 percent in 2000, and 33 percent in 2001.[6] The year-to-year changes are small and do not reveal a trend, but these IMA surveys suggest that while performance measures are changing

© Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
0521812437 - Rethinking Performance Measurement: Beyond the Balanced Scorecard
Marshall W. Meyer
Excerpt
More information

rapidly, management accountants do not experience these changes as improvements.

Avoiding bedrock issues: the "balanced scorecard"

Firms and non-business organizations alike can no longer afford to avoid bedrock issues of performance measurement. Let's be frank. For the last decade, discussion of performance measurement has been dominated by the "balanced scorecard." Many books, articles, and cases about the balanced scorecard have appeared during that period, the *Harvard Business Review* has called the balanced scorecard one of the most important management ideas in the last seventy-five years, and an organization called the Balanced Scorecard Collaborative serves as a central clearing house for what it calls the "balanced scorecard movement."[7] What is missing from the spin surrounding the balanced scorecard is a simple fact about performance measures, the significance of which is not widely appreciated: common-sense measures used to gauge the performance of a firm are generally uncorrelated. In other words, look across a large number of firms or their business units and you will find that profitability, market share, customer satisfaction, and operating efficiency are weakly and sometimes negatively correlated. These measures move in different directions about as often as they move in tandem. Social scientists have known this for years and have drawn two conclusions. First, measuring performance is difficult (since it is not clear that performance is a single construct). Second, the choice of performance measures is often arbitrary (since it is difficult to prove that any one measure is better than others). Though neither of these conclusions is particularly useful, they would not surprise managers.

Beginning in 1992, Robert Kaplan and David Norton transformed the persistent observation that measures are generally uncorrelated into a prescription for business practice: just as pilots track multiple instruments to gauge the performance of an aircraft, managers should track multiple measures to gauge the performance of their firms. "Managers want a balanced presentation of both financial and operational measures . . . The scorecard brings together, in a single management report, many of the seemingly disparate elements of a company's competitive agenda . . ."[8] Not only is the analogy between cockpit instruments and the measures needed to guide firms compelling, but its logic is also

© Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
0521812437 - Rethinking Performance Measurement: Beyond the Balanced Scorecard
Marshall W. Meyer
Excerpt
More information

impeccable. Consider the counterfactual. Ask whether multiple measures would be necessary if measures were strongly correlated, that is if the most common performance measures rose and fell together. The answer is this: if performance measures were strongly correlated, then all would contain essentially the same information, any one of them would contain complete information about the performance of the firm, and there would be no need for multiple measures or a "balanced scorecard."[9] For example, if customer satisfaction and bottom-line results were strongly correlated, there would be no need, except for comfort, to measure customer satisfaction since bottom-line results would signal the level of customer satisfaction. Now consider the actual. Again, performance measures are weakly correlated. Each contains different information about the performance of the firm, and scorecards utilizing multiple measures are needed to capture the performance of the firm completely. In other words, customer satisfaction (and operational performance, innovation, and so on) must be measured alongside financial results *because they are different*.

Unfortunately, the logic lying behind the scorecard approach to performance measurement can go awry when measures are put to use. While there are good reasons to measure multiple dimensions of performance, there are also strong pressures to appraise performance along one dimension: better or worse. These pressures are strongest when compensating and rewarding people's performance, but they are also present when making investment decisions. Whenever managers ask whether firm A performs better than B, whether division C performs better than D, or, most poignantly, if employee E is a better performer and hence should be compensated more generously than F, G, and H, they are tacitly if not explicitly trying to reduce performance to a single dimension.

Even Kaplan and Norton recognize these limitations of the "balanced scorecard" and are reluctant to recommend scorecards to appraise and compensate performance. Consider the following:

> Norton: . . . firms often hesitate to link the scorecard to compensation. Kaplan: They should hesitate, because they have to be sure they have the right measures [on the scorecard]. They want to run with the measures for several months, even up to a year, before saying they have confidence in them. Second, they may want to be sure of the hardness of the data, particularly since some of the balanced scorecard measures are more subjective. Compensation is such a powerful lever that you have to be pretty confident that

© Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
0521812437 - Rethinking Performance Measurement: Beyond the Balanced Scorecard
Marshall W. Meyer
Excerpt
More information

you have the right measures and have good data for the measures [before making the link].[10]

Note that Kaplan and Norton construe the compensation problem narrowly, as a problem of finding the "right measures." The compensation problem, in fact, is much broader. It exposes the tension between measuring performance along several dimensions and appraising performance ultimately on one dimension. Remember: scorecard measures are necessarily different. If they weren't, then they would be redundant and there would be no need for the balanced scorecard because any one measure would do. The compensation problem, moreover, raises the question of whether the "right measures" can in fact be found. "Right measures," to be sure, can be found in static environments where the parameters of performance are well understood. Go back to the cockpit analogy. Pilots know how an aircraft must perform in order to complete its mission and rely on their instruments to compare actual to required performance. In competitive environments, however, the performance required to produce a satisfactory return can change unpredictably; in other words, measures that were right can be rendered obsolete or pernicious overnight.

Rather than tackling these bedrock problems of performance measurement, Kaplan and Norton have recast the "balanced scorecard" as a management system intended to communicate strategies and objectives more effectively than non-scorecard systems: "Measurement creates focus for the future. The measures chosen by managers communicate important messages to all organizational units and employees . . . the Balanced Scorecard concept evolved from a performance measurement system to become the organizing framework, the operating system, for a new strategic management system."[11] I am skeptical about basing strategy on performance measures. I worry about unintended consequences, especially unintended consequences of imperfect measures – as will be shown, all performance measures are imperfect. In particular, I worry about measurement systems becoming arteriosclerotic, turning into the rigid quota systems that ruined socialist economies. "What you measure is what you get" captures the problem: if you cannot measure what you want, then you will not get what you want.

I'm not saying that we can do without performance measures, but I am saying that we should tackle bedrock issues before basing strategies

© Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
0521812437 - Rethinking Performance Measurement: Beyond the Balanced Scorecard
Marshall W. Meyer
Excerpt
More information

on such measures. Again, the specter of quotas haunts me. I think that we should approach the bedrock issues realistically. We should assume that measuring performance is difficult. If performance measurement weren't difficult, then it wouldn't be the chronic problem that it is. I also think we should assume that performance measurement is difficult for good reasons. The good reasons, I suspect, lie in both the nature of organizations and the people in them.

Consider organizations first. The dilemma created by organizations is illustrated by Adam Smith's pin-making factory, where every worker is like an independent business – one cuts wire, a second sharpens the wire, a third solders pin heads onto the sharpened wire, a fourth boxes pins, and so forth – engaging in cash transactions with co-workers. There is no performance measurement problem because each worker has his or her own revenues and costs. There is an efficiency problem, however, since intermediate inventories will accumulate if workers fail to coordinate their efforts and produce at different rates – if the wire cutter works faster than the sharpener, for example. The solution to the efficiency problem is placing the workers under a common supervisor charged with coordinating the process; in other words, creating an organization. But solving the efficiency problem creates a performance measurement problem. There is no simple way to measure separately the contributions of the wire cutter, the wire sharpener, the solderer, and the boxer to the performance of the organization that has been created because one revenue stream has replaced the independent revenue streams that formerly existed.

Now consider the people problem. People will assume performance measures to be consequential and will strive to improve measured performance even if the performance that is measured is not the performance that is actually sought – teaching to test is illustrative. Performance measures, as a consequence, get progressively worse with use, and managers face the challenge of searching out newer and better measures – better, that is, until they deteriorate – while retaining the semblance of clarity and consistency of direction. That organizations and the people in them create impediments to measuring performance as well as we would like is central to the rethinking of performance measurement I shall propose.

The message and metaphor of the balanced scorecard were, of course, important first steps in getting at bedrock issues of performance measurement. The notion that a tool as complicated as a baseball

© Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
0521812437 - Rethinking Performance Measurement: Beyond the Balanced Scorecard
Marshall W. Meyer
Excerpt
More information

scorecard might be needed to gauge corporate performance has jarred managers into realizing there is more to performance than the bottom line. But the message and the metaphor are now ten years old. It is time to rethink performance measurement once more.

Ideal performance measurement

The rethinking of performance measurement begins with a simple question: what properties do we look for in performance measures? Ideally, the performance measures of choice would meet the following requirements:

- Parsimony. There would be relatively few measures to keep track of, perhaps as few as three financial measures and three non-financial measures. (I have chosen three plus three arbitrarily, but I think these numbers are realistic.) Cognitive limits would be exceeded and information would actually be lost were there many more measures.
- Predictive ability. The non-financial measures would predict subsequent financial performance, in other words, the non-financials would serve as leading performance indicators and the financials as lagging indicators, as measures summarizing performance after it occurred. Non-financial measures not demonstrated to be leading indicators would be discarded unless, of course, they were tracked as matters of regulation, ethics, and security – "must-dos" for firms.
- Pervasiveness. These measures would pervade the organization – the same measures would apply everywhere. Measures pervading the organization have three key advantages over highly specific measures: they can be summed from the bottom to the top of the organization, which allows people to see connections between their results and the results of the firm; they can be decomposed downward, which gives senior managers drill-down capability; and they can be compared horizontally across different units, which facilitates improvement and performance appraisal.
- Stability. The measurement system would be stable. Measures would change gradually so as to maintain people's awareness of long-term goals and consistency in their behavior.
- Applicability to compensation. People would be compensated for performance on these measures, that is for financial results and results of non-financial measures known to be leading indicators of financial results.

© Cambridge University Press www.cambridge.org

The requirements of ideal performance measurement are very stringent, far more stringent than the requirements of the balanced scorecard. The balanced scorecard imposes only the two requirements on measures, parsimony and predictive ability: in principle, scorecard measures are more parsimonious than the potpourri of measures tracked by most large firms, and non-financial scorecard measures predict financial results. The scorecard does not address pervasiveness other than acknowledging that scorecards and scorecard measures are likely to vary across different parts of the organization. Nor does the scorecard address the stability of measures. Moreover, as noted, Kaplan and Norton are cautious about using scorecard measures to compensate people – for good reason, as will be seen below.

Rarely if ever do we find performance measures meeting these common-sense requirements. Here is why:

- Firms are swamped with measures, and the problem of too many measures is, if anything, getting worse, the balanced scorecard withstanding. It is commonplace for firms to have fifty to sixty top-level measures, both financial and non-financial. One of the longest lists of top-level measures I have seen includes twenty financial measures, twenty-two customer measures, sixteen measures of internal process, nineteen measures of renewal and development, and thirteen human resources measures.[12] Many firms, I am sure, have even more top-level measures.
- Our ability to create and disseminate measures has outpaced, at least for now, our ability to separate the few non-financial measures containing information about future financial performance from the many that do not. To be sure, research studies show that a myriad of non-financial measures such as customer and employee satisfaction affect financial performance, but their impact is modest, often firm- and industry-specific, and discoverable only after the fact.
- Few non-financial measures pervade the organization. It is easier to find financial measures that pervade the organization, but keep in mind that many firms have struggled unsuccessfully to drive measures of shareholder value from the top to the bottom of the organization.
- Performance measures, non-financial measures especially, never stand still. With use they lose variance, sometimes rapidly, and hence the capacity to discriminate good from bad performance. This is the use-it-and-lose-it principle in performance measurement. Managers respond by continually shuffling measures.

© Cambridge University Press

CAMBRIDGE

Cambridge University Press
0521812437 - Rethinking Performance Measurement: Beyond the Balanced Scorecard
Marshall W. Meyer
Excerpt
More information

- Compensating people for performance on multiple measures is extremely difficult. Paying people on a single measure creates enough dysfunctions. Paying them on many measures creates more. The problem is combining dissimilar measures into an overall evaluation of performance and hence compensation. If measures are combined formulaically, people will game the formula. If measures are combined subjectively, people will not understand the connection between measured performance and their compensation.

There is a still more fundamental reason for the gap between ideal performance measurement and performance measurement as it is. The modern conception of performance, which is the economic conception of performance, renders the performance of the firm not entirely measurable. The modern conception of performance is future cash flows – "cash flows still to come"[13] – discounted to present value. In other words, we think of the firm as assets capable of generating current and future cash flows.[14] Future cash flows, by definition, cannot be measured. Nor can we measure the long-term viability and efficiency of the firm in the absence of which cash flows will dwindle or vanish. What we can and do measure are current cash flows (financial performance), potential predictors of future cash flows (non-financial measures), and proxies for future cash flows (share prices). All of these are imperfect. They are, at best, second-best measures. Note the paradox that is at the heart of efforts to improve performance measurement: knowing that most measures are second best compels us to search for better measures that are inevitably second best. If we had a different conception of performance – for example if we believed a firm's performance was its current assets rather than future cash flows – then measuring the performance of the firm would be no more complicated than measuring the performance of an airplane. One point deserves emphasis: I'm not saying that everyone subscribes to the notion of economic performance, of performance as future cash flows or even as the long-term viability and efficiency of the firm. Managers, in particular, think of performance as meeting the targets they have been assigned. I *am* saying, however, that our unease with most of the performance measures we have is due to the gap between what we can measure – current financial and non-financial results – and the future cash flows we would measure if we could.

© Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
0521812437 - Rethinking Performance Measurement: Beyond the Balanced Scorecard
Marshall W. Meyer
Excerpt
More information

The performance chain

To search intelligently for better, albeit second-best, performance measures, we may have to rethink the firm and the relevant units for measuring performance. Right now, we think of firms as black boxes: investment flows into the firm, activities take place, products are made and sold to customers as a result of these activities, and an income statement, balance sheet, and market valuation of the firm follow. Since financial results – the income statement, balance sheet, and market valuation – accrue to the firm as a whole or, internally, to large chunks of the firm called business units, we look for drivers of financial performance, that is non-financial measures describing internal processes, products, and customers, at the level of the entire firm or its business units. The problem with the black-box approach to the firm and performance measurement is that it masks differences within firms and their business units: so many processes take place, so many products are produced, and so many customers are served that firm- or business unit-level performance measures – which I'll call aggregate measures – conceal important sources of variation. The things a firm does well are lumped together with the things it does poorly, making it difficult to know, for example, precisely where to invest and where to cut costs. Importantly, the larger the firm and its business units, the more information about performance is obscured by aggregate performance measures.[15]

The rethinking of the firm and of the relevant units for measuring performance begins by asking where the performance of the firm comes from. The performance of the firm originates in what the firm does, in its activities or routines. These activities give rise to costs, but they also generate revenues in excess of costs to the extent that the firm's products and services add value for customers. These cash flows and the expectation of future cash flows in turn give rise to the valuation of the firm in capital markets. The causal chain running from activities to costs to revenues to the valuation of the firm in capital markets is shown in figure I.1. This 'performance chain' is an extension of Michael Porter's idea of the value chain that incorporates costs.[16]

The performance chain carries some immediate implications for performance measurement. First, the units in the performance chain bear little resemblance to the units on a typical organization chart. There are three principal units: the firm, the customer, and the activity. By

© Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
0521812437 - Rethinking Performance Measurement: Beyond the Balanced Scorecard
Marshall W. Meyer
Excerpt
More information

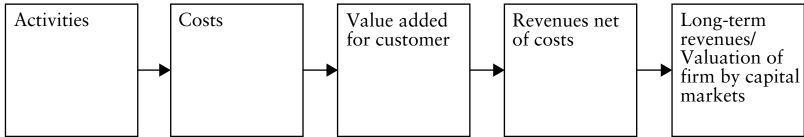

Figure I.1 The performance chain of the firm.

contrast, the units displayed on an organization chart are typically the firm, business units, functional units, and work groups within business and functional units. Many activities take place within business units, functional units, and work groups, and many customers are served, directly or indirectly, by each of them. The performance chain thus raises two questions: should firms be partitioned into units, such as activities, that are much smaller than the units shown on organization charts, and how should performance be measured on these smaller units?

Second, the performance chain shows that activities incur costs and customers supply revenues – and that revenues and costs are usually joined at the level of the firm. This raises the question of whether costs can be assigned to customers and, correspondingly, whether revenues can be assigned to activities so that revenues and costs can be compared for individual customers and activities. It is not uncommon for firms to assign costs to customers and then compare revenues to costs customer by customer. This is sometimes called customer profitability analysis. I will show below that once you assign costs to customers, you can also assign revenues to activities, in other words, you also can also compare revenues to costs activity by activity. I call this activity-based profitability analysis or ABPA. The possibility of assigning revenues and costs to individual customers and activities is one of several reasons why it may be better for performance measures to follow the performance chain than to follow the organization chart – while you can always assign costs to the units shown on an organization chart, you cannot easily assign revenues to units smaller than your profit centers or strategic business units.

The elemental conception of the firm

The performance chain also carries implications for how we think about the firm itself. Put aside your preconceptions about organizations and imagine the firm as a bundle of activities, nothing more.

© Cambridge University Press www.cambridge.org