

PREVENTING BIAS IN THE BALANCED SCORECARD IMPLEMENTATION FOR PERFORMANCE ASSESSMENT

EKA ARDHANI SISDYANI

Jurusan Akuntansi
Fakultas Ekonomi Universitas Udayana

Abstract

Balanced scorecard has been widely recognized as a tool of comprehensive performance measurement both for small and large organizations. However, when the method is employed to compare performance of two or more business units, evaluator might encounter with common-measures bias. This bias would emerge when evaluator only pay attention to common measures existed in the units being evaluated, and ignore the unique measures belong to each individual unit. This condition would mislead evaluator to reach inaccurate conclusion regarding the achievement of the units, and consequently could end up with incorrect decision, such as compensation policy. Such a bias can be reduced by employing a technique that has been empirically studied by Roberts, Albright, and Hibbets (2004), called disaggregated plus mechanically aggregated. This technique has successfully reduced common-measures bias presented in the absence of a treatment using this technique.

Keywords : balanced scorecard, common-measure bias, performance

I. BACKGROUND

The concept of balanced scorecard (BSC) has been introduced for more than a decade by Kaplan and Norton in 1992. In the early days of implementation, BSC was used as a performance measurement method, but then its functions widen to be as a tool for implementing and monitoring strategy of an organization (See Kaplan and Norton, 1996; Mulyadi, 2005). Since its inception in 1992, BSC has been widely used by both large and small companies (Silk 1998; Frigo and Krumwiede 2000). The proponents of BSC argue that BSC provide a multidimensional view of organizational performance, and

thus enable managers to utilize strategically important non-financial as well as financial measures.

To be able to adopt the BSC, each business unit in a firm must develop its own scorecard with measures that capture not only common strategy of the unit but also its unique strategy. Common measures are applied to every business unit in the firm; on the other hand, unique measures are only applied to a specific unit. Each unit would have different unique measures depending on its specific environment. In the process of measuring performance of organization, especially when the performance comparison among business units is made, for example to link the result to compensation scheme, it is important to take into account not only the common measures, but also the unique measures of the units.

Lipe and Salterio (2000) identified a potential problem in using BSC to evaluate performance. Their empirical study resulted that when evaluating divisional performance, the M.B.A. students (subjects of their research) who are assigned the role of superior managers ignored unique BSC measures. Superiors relied only on the items appearing on both divisions' scorecard. Half of the measures included in the scorecards, which were unique or specific to a single division, were ignored. In other words, they found a domination of common measures which they call *common-measures bias*.

The study by Lipe and Salterio (2000) then was extended by Dilla and Steinbart (2005). Dilla and Steinbart (2005) provided their experiment's participants a sufficient training and practical sessions of designing and implementing BSC as a tool of performance measurement before assigning them to complete the task in the research design. They used participants who have more than a cursory understanding of the BSC.

The participants learned about BSC through lectures and readings and by developing actual BSCs for two different organizations. So the participants in Dilla and Steinbart's study were more knowledgeable and experienced in BSC than participants of Lipe and Salterio.

The study of Dilla and Steinbart (2005) showed that more knowledgeable and experienced decision makers would use both common and unique BSC measures to evaluate subordinates' performance and consequently to allocate bonuses, but still, placed greater emphasis on common measures compared to unique measures when making both judgments. This finding shows that although common-measures bias decreases, it is still in existence. It then would imply that managers may not give much attention to factors that they perceive as not affecting their compensation. If unique measures reflect key factors of a unit's strategy, then inattention to them will undermine the usefulness of the BSC as a strategic management system (Dilla and Steinbart 2005). Therefore it is necessary to find a way to eliminate the bias due to unequal attention for common versus unique measures in performance evaluation process.

II. Preventing Common-Measures Bias in Balanced Scorecard for Performance

Evaluation

Lipe and Salterio (2000) state that the common measures employed in the BSC tend to be more traditional financial measures and tend to lag actual performance. Examples of common measures in financial perspective would be return on sales, sales growth, and return on assets. In contrast, unique measures that include sales of new stores, revenue per sales visit, and catalog profits, tend to be nontraditional and more

importantly, leading indicators of performance. The unique measures as the leading indicators will capture elements of corporate and division strategic emphasis, thus it would be problematic to ignore them.

Common measures may dominate in comparative evaluations for at least three related reasons. First, they form a smaller subset of the total information, and it is cognitively easier to retain and process less, rather than more, information (Anderson 1990). Second, not only does this result in less total information, but also it may result in fewer categories or types of information to process (Lipe and Salterio 2002). Third, common measures are the only information available to directly compare the managers (Roberts, Albright, and Hibbets 2004).

According to Lipe and Salterio (2000) their experiment participants ignored unique measures in order to reduce their efforts to complete the evaluation tasks. Kennedy (1995) suggested that one way to improve decision when efforts are not sufficient is to use a decision aid, such as statistical modeling combined with human judgment. When this approach is applied to the balanced scorecard (BSC), Roberts, Albright, and Hibbets (2004) proposed a two-step process to be followed. First, evaluation decision is disaggregated into several smaller decisions; and second, these smaller decisions then are aggregated into an overall score based on the weight that is predetermined by the decision maker. The process will be explored further in the following paragraphs.

Step one is done by disaggregating a complex decision which would encourage the extent to which each individual dimension is processed. If a decision maker just focus attention to one limited dimension, his/her short term working memory would be free

from simultaneously keeping information about other dimensions. This process must be done for each dimension being evaluated one by one. This shift in attention and processing capacity should facilitate greater total effort and ensure that effort is exerted on all measures. This step would overcome common-measure bias to the extent the bias is caused by failure to adequately attend to the unique measures.

In step two, all dimensions that have been evaluated individually in step one, now are aggregated into an overall score by utilizing the predetermined weights. The weights should reinforce the importance of both common and unique measures to the organization. Therefore it is more likely that both measures will be used in subsequent holistic evaluation because the decision maker will have already finished the process of evaluating both measures when he/she evaluate each dimension in step 1.

To implement the procedures, first the decision maker has to rate each manager's performance on each of the performance indicators in BSC both common and unique measures, for instance, using a scale from 0 (unacceptable) to 100 (excellent). After that, the evaluator multiplies these individual judgments by predetermined weights and sums the weighted scores to end up with a total aggregated score for each manager.

The strategy of disaggregating decision would be more useful when the decision is more complex (Roberts, Albright, and Hibbets 2004). In BSC performance evaluation, normally there will be four to seven performance measures in each of four categories (financial, customer, internal business process, and learning and growth), as proposed by Kaplan and Norton (1996). Consequently, all in all evaluators using BSC method could potentially have 16 to 28 indicators to take into account in assessing the performance of a firm manager. Thus, judging a performance with BSC is complex enough to realize the

benefits of disaggregated plus mechanically aggregated approach. This approach would decrease cognitive demands at any one time because the amount of information to be considered for evaluating each individual dimension is less than the information in the entire BSC. However, this approach would also have a drawback, that is there will be more time and effort needed because the number of evaluation and computations increase. Alternatively, performance can be evaluated for each of the four perspective of BSC at a time, and continued by making a holistic judgment. This would enable the evaluator to substantially lessen the amount of information to be processed at each stage, and thus reduce the number of evaluation and computation packages.

The use of disaggregated plus mechanically aggregated approach is not relevant when performance assessment is done for an individual manager or firm, because in this case common-measures bias will not exist (Slovic and MacPhilyamy 1974 as stated in Roberts, Albright, and Hibbets 2004).

III. Concluding Remarks

Comparing performances of two or more divisions or firms must be done carefully. Utilizing the balanced scorecard (BSC) approach to assess and compare managers' performance would need a caution of the existence of common-measures bias, because evaluator tends to only consider the common measures and conclude the performance based on this type of measures. In fact, some unique measures are also important to take into account, because performance is built up from common as well as unique characteristics of a business unit. Evaluator who ignores unique measures may

end up with inaccurate judgment. Therefore it is important to eliminate the common-measures bias in the performance appraisal process using BSC method.

One way to do so is by applying a technique called disaggregated plus mechanically aggregated as suggested by Roberts, Albright, and Hibbets (2004). This technique has been empirically studied and successfully demonstrated that common-measures bias could be reduced significantly.

REFERENCE

- Dilla, W. N., and P. J. Steinbart. 2005. "Relative Weighting of Common and Unique Balanced Scorecard Measures by Knowledgeable Decision Makers". *Behavioral Research in Accounting* Vol.17: 43—53.
- Friego, M. L., and K. R. Krumwiede. 2000. "The Balanced Scorecard: A Winning Performance Measurement System". *Strategic Finance*. January: 50—54.
- Kaplan, R. S., and D. P. Norton. 1992. "The Balanced Scorecard: Measures that Drive Performance". *Harvard Business Review*. January-February: 71—79.
- Kaplan, R. S., and D. P. Norton. 1996. *The Balanced Scorecard*. Boston: Harvard Business School Press.
- Lipe, M., and S. Salterio. 2000. "The Balanced Scorecard: Judgmental Effects of Common and Unique Performance Measures". *The Accounting Review* 75 (3): 283—298
- Mulyadi. 2005. *Sistem Manajemen Strategik Berbasis Balanced Scorecard*. Yogyakarta: UPP AMP YKPN.
- Roberts, M. L., T. L. Albright, and A. R. Hibbets. 2004. "Debiasing Balanced Scorecard Evaluations". *Behavioral Research in Accounting* Vol.16: 75—88.
- Silk, S. 1998. "Automating the Balanced Scorecard". *Management Accounting* (May): 38—44.