SUMMER 2013 VOL.54 NO.4

MITSloan
Management Review

Analyzing Performance in Service Organizations

By H. David Sherman and Joe Zhu

REPRINT NUMBER 54413

Analyzing Performance in Service Organizations

A technique called balanced benchmarking provides managers with a sophisticated mechanism by which to assess and manage the effectiveness of different branches or units.

BY H. DAVID SHERMAN AND JOE ZHU

THE LEADING QUESTION

How can service businesses measure and manage productivity?

FINDINGS

- Balanced benchmarking is a linear programming technique that enables organizations to compare different service providers.
- Companies can use balanced benchmarking to identify both best practices and inefficiencies.
- The technique also allows managers to test their assumptions about productivity.

IN HIS 2003 BOOK *Moneyball: The Art of Winning an Unfair Game*, Michael Lewis described how the Oakland Athletics baseball team used statistical analysis to identify undervalued players.[1] One lesson from the baseball world of "moneyball" is that we can't always trust our intuition about how employees will perform. Savvy business managers know that their intuition can often be misleading, if not downright incorrect. And just as sports teams have increasingly relied on rigorous quantitative analyses, so have many businesses.

In particular, a growing number of service businesses have been investigating the use of a sophisticated linear programming technique called DEA, or data envelopment analysis. (In this article, we use the term "balanced benchmarking" to denote DEA.) The technique enables companies to benchmark and locate best practices that are not visible through other commonly used management methodologies. (See "The Basics of Balanced Benchmarking," p. 38.)

When it was first introduced in the 1980s,[2] balanced benchmarking was an academic tool for measuring and managing the relative efficiency of peer organizations. Balanced benchmarking required the adaptation of various computer programs, so its use in the 1980s was limited to a small group of academics and practitioners with linear programming expertise. Early users were able to apply and generate results from balanced benchmarking that demonstrated its effectiveness, but its inaccessibility limited its independent adoption and application by managers. However, shortly after 2000, balanced-benchmarking algorithms were adapted for

Excel software — making it accessible to users with little or no knowledge of linear programming.[3]

Balanced benchmarking is unique both in its ability to identify paths to improve productivity and in its value as a complement to other analytic techniques. Balanced benchmarking simultaneously considers the multiple resources used to generate multiple services, along with the quality of the services provided. For example, bank branches can use six or more types of resources and provide 20 or more types of services, all of which are considered with balanced benchmarking. By combining this information, balanced benchmarking provides unique insights about best practices and opportunities to improve productivity and profitability — information not available with other techniques.

Various studies have investigated how balanced benchmarking improves the performance of service organizations, including banks, insurance companies, hotels, real estate agents, customer service representatives, computer manufacturer field service organizations, auto dealers, health-care organizations and supply chains. In the nonprofit arena, the methodology has been used to investigate the efficiency of government agencies, school systems and universities. Some of the more interesting applications from around the world have included stores from a major Fortune 500 multinational retail chain,[4] fish farms in China,[5] hedge fund performance,[6] power plants in Israel[7] and so on.

THE BASICS OF BALANCED BENCHMARKING

Balanced benchmarking is a linear programming technique that was originally developed to evaluate nonprofit and governmental organizations, but it has subsequently been applied to the service operations of a variety of private companies. One of the many advantages of balanced benchmarking (originally called data envelopment analysis) is that it allows a company to compare various business units (for example, the different stores of a national chain) in terms of different inputs (the number of sales clerks and managers, the square footage of the display space, the inventory, the advertising expenditures, the utilities used and so on) that are used to generate a number of outputs (total revenues, profits, number of customers served, average sales and number of items purchased per customer, customer satisfaction ratings and so on).

For simplicity, we consider the hypothetical example of a small chain of five custom-tailor shops, and we take into account just two inputs: employees (measured in labor hours, or *H*) and supplies (measured in dollars, or *S*). We further assume that all five shops have the same daily revenues and services, but their use of inputs to generate that output varies widely.

From this simple example, it's readily apparent that store *B1* is not as efficient as *B4* because it uses 100 more units of S (dollars) and the same amount of H (hours) to achieve the same output, and store *B2* is not as efficient as *B4* because it uses 10 more units of *H* and the same amount of S to achieve the same output. But such observations aren't always so obvious if a company has hundreds of branches and is considering more than a dozen inputs and several outputs. In such situations, identifying which business units are relatively inefficient is hardly straightforward.

STORE BRANCH	DAILY INPUTS USED	DAILY OUTPUT GENERATED
B1	20 units of *H* 300 units of *S*	$1,000 in sales
B2	30 units of *H* 200 units of *S*	$1,000
B3	40 units of *H* 100 units of *S*	$1,000
B4	20 units of *H* 200 units of *S*	$1,000
B5	10 units of *H* 400 units of *S*	$1,000

We should note that, although the analysis can identify stores *B1* and *B2* as being relatively inefficient, it can't determine which of the other stores (*B3*, *B4*, or *B5*) is the most efficient, because each uses a different balance of inputs that can't definitively be compared. For example, *B3* uses more *H* and less *S* than *B2*, which may reflect different methods and styles of providing services, and one is not clearly superior based on these data. *B3*, *B4*, and *B5* are all then considered relative best-practice shops because no other shop in the analysis is definitively more efficient than these shops. So, generally speaking, the inefficient branches aren't held up to the standards of one ideal store but rather to a reference set of other stores.

Based on a balanced-benchmarking analysis, *B1* would have an efficiency rating of 85.7%. This means that, based on a comparison with a group of relative best-practice shops, *B1* could theoretically use just 85.7%[i] of its current input to generate the same level of output.[ii] In other words, *B1* could cut back to just 17 employee hours (0.857 × 20) and 257 supply dollars (0.857 × 300). Of course, other factors could legitimately be contributing to the relative inefficiency of store *B1*, but, at the very least, balanced benchmarking helps focus management's attention on those locations that might warrant further investigation. Furthermore, this analysis steers management's attention to the most efficient business units — namely, the benchmark shops that might be a source of best practices that could be transferred to other locations.

The power of balanced benchmarking is, of course, more apparent when there are many service units being evaluated, up to, say, 3,000 physicians or 2,000 bank branches, where the units are using multiple resources to provide several services (sometimes more than 20 types of service) and where these units might be located near each other or in different states or different countries.

The benefits of balanced benchmarking are numerous. First and foremost, managers don't need to fly blind. This is especially valuable in service businesses, which present unique managerial challenges. For example, while the quality of a manufactured product can be tested and inspected prior to putting the product on the shelf for sale, the quality of services is dependent on the services provider at the time the service is delivered. The production process and cost are influenced by both the provider and customer, and the complexity of that interaction can exceed the most complicated manufacturing activities, particularly in professional services such as health care and management consulting.

Balanced benchmarking provides managers with a sophisticated mechanism to assess the performance of different service providers — comparing, for example, the London and Tokyo offices of a global advertising agency — by going well beyond crude metrics and ratios such as profitability and account billings per employee. From the results of balanced benchmarking, a company can identify its least efficient offices or business units, and it can assess the magnitude of the inefficiency and investigate potential paths for improvement that the analysis has identified. Moreover, executives can study the top-performing units, identify the best practices and transfer that valuable knowledge throughout the organization to enhance performance. Lastly, balanced benchmarking enables companies to test their assumptions, particularly before implementing cost-cutting initiatives that might inadvertently be counterproductive.

Understanding Balanced Benchmarking

The calculations involved in balanced benchmarking are intensive, but the overall approach is straightforward. The technique essentially looks at what inputs (various resources, including labor) are being used to produce which outputs (the services provided). It then compares different business units — for example, the various stores of a large retail chain — based on their input levels, output levels and quality measures and identifies which of them are the most and least efficient.

To gain a deeper understanding of the balanced benchmarking technique, consider the following example. A U.S. bank with more than 200 branches in five states wanted to reduce its operating costs but had no general benchmarks for doing so. The one tool that the bank used was a staffing model based on total teller transactions and peak demand periods. But this model was just for tellers, who handle basic transactions like deposits and withdrawals. Branches typically also include managers and platform personnel, who are responsible for more complicated dealings such as loan applications, online checking and the opening of individual retirement accounts. Moreover, the bank had information about most of its customer transactions (branches might provide more than 20 different types of transactions), but those data weren't being analyzed or used to evaluate the efficiency of its operations. The bottom line was that executives didn't really know exactly how efficiently each branch was operating — whether, for example, a particular site had the right ratio of tellers, platform personnel and managers.

So the bank conducted a balanced benchmarking analysis of its branches. The study considered several inputs, including the number of tellers, platform personnel and managers, in addition to various costs for supplies, local advertising, telecommunications and travel. And it looked at the outputs of each branch in terms of various transactions such as deposits, withdrawals, checks cashed, safe-deposit visits, new accounts opened, mortgage and consumer loans processed and so on. To assess service quality, companies have used a variety of methods, including customer surveys and questionnaires. The bank decided to rely on evaluations by "mystery shoppers" posing as customers, because management felt that technique captured the data most relevant to the analysis.

Of the total number of branches, 46 were placed in the benchmark reference set. These branches had both high-quality service and high efficiency (that is, their output was relatively high with respect to their input utilization). Another 32 branches were identified as highly efficient but with low quality. These branches were not allowed to serve as benchmarks until their quality level could be improved to a minimum threshold. Of the remaining branches, 147 had efficiency ratings at or below 90%, including 42 that had a rating below 60%. (A branch with

a rating of 60% suggests that it might be using up to 40% more resources than the best benchmark branches to provide the same volume, mix and quality of services.)

Regional bank executives met with each local manager of a branch whose efficiency rating was less than 90%. The goal was to identify any opportunities for cost savings. As just one example, the analysis revealed high telephone charges in two states. One of those states contained more than 30 branches, and management was able to negotiate new phone contracts for significant cost savings in that area. Furthermore, the overall analysis revealed that the bank could theoretically cut branch staff by 21% without any drop in the output of work or its quality. Of course, such theoretical savings aren't always achievable in practice. One group of branches, for example, resisted reducing its staff by the suggested number of 60 full-time employees. Instead, the actual reduction was just six full-time individuals after the regional manager argued for retaining more staff because of changing market conditions and the need to build business at those branches. As it turned out, the bank was able to reduce total branch staff by 7.4% within six months of completing the analysis, far short of the theoretical 21% but still a substantial cost savings. In addition, the analysis helped identify other areas of potential cost savings that might be investigated.

More importantly, balanced benchmarking helped the bank avoid making a major mistake. Previously, executives had been considering closing smaller branches located in vacation spots, retirement communities and low-income urban neighborhoods. The assumption was that these types of branches couldn't achieve maximum efficiency because of seasonal staffing requirements, slower transaction times and a higher prevalence of multilingual customers. Moreover, smaller branches were typically thought to be less efficient because they need a minimum staff level to maintain adequate financial controls over certain transactions that might occur only occasionally, whereas large branches can enjoy greater economies of scale. But the balanced benchmarking study found that some of the smaller branches were among the best performers, while many of the largest branches were found to be inefficient. Indeed, management learned that large branches with high deposits might appear to be very profitable when in fact they could be using a significant number of excess personnel.

Management also looked at the benchmark branches to identify any best practices that might help increase the efficiency of the organization as a whole. One practice identified was the aggressive use of part-time employees to better match staff capacity with work demands. Because many of the low-performing branches had trouble attracting and hiring part-timers, the bank changed its policies to provide better health plans and other benefits to its part-time employees.

Balanced Benchmarking Lessons for Your Business

In addition to our work with the aforementioned bank and several other U.S. banks, we have helped various other organizations implement balanced benchmarking to improve their operations, and we have reviewed dozens of studies in numerous industries. From that research, we have culled a number of managerial lessons for companies to get the most out of applying the technique.

Even when the desired data are scarce, efficiency can still be assessed. Many companies are awash in data. Some retailers, for example, collect copious real-time information of exactly what sells when. This, however, doesn't necessarily mean that managers will always have the data they need or desire, but balanced benchmarking can often be performed using information that is readily available. When determining the inputs and outputs to be used in any analysis, companies can ensure "buy-in" by involving the managers of the business units in the process of identifying and incorporating all the relevant resources used and services provided by a business unit. This will help minimize any "push back" should the analysis yield some unflattering results.

Don't prescreen. Some companies make the mistake of screening out business units that they think are outliers because they don't want to bias or corrupt their results. But often those "outliers" contain information that is important to the analysis. Those business units might, for example, be deploying a best practice that other groups could

benefit from adopting. Of course, some business units should be screened out: A retailer might, for example, omit new stores that don't have enough of a track record. But companies should nevertheless be careful about prematurely screening out sources of information that could be invaluable.

Look for major clusters. In the initial analysis, managers should look for major clusters in the results. Often, for example, larger business units will form a cluster, indicating that efficiency of scale is a major factor. Or the cluster might indicate the effects of a major policy difference. A balanced-benchmarking analysis of professional sports teams in the United States,[8] for example, found that franchises in the National Football League tend to be more efficient than those in Major League Baseball because of the NFL's policies on revenue sharing and salary caps. Major clusters might also indicate environmental or structural factors. Consider a study of medical centers for veterans in the United States,[9] which found that an important differentiator was whether a center was affiliated with a university. Those that were affiliated generally had lower efficiency, presumably because the case mix tends to be more complex in such hospitals, thus requiring more labor (physicians, nurses and other staff), equipment and medical supplies, including drugs.

Identify best practices. Once major clusters are identified, the balanced benchmarking analysis can be run again. In the study of medical centers for veterans, for example, a subsequent round of balanced benchmarking focused on 120 hospitals with university affiliations. From that iteration, managers can identify the best and worst performers of that particular large cluster and investigate what might be causing the difference. Of particular importance are any best practices that could be transferred to other locations. A discount brokerage company, for instance, found that at a top-performing branch the manager had cross-trained employees so that workers could fill in for other functions when needed. This insight generated changes in recommended training practices at other offices as well as consideration of alternate physical layouts to encourage wider use of these practices.

Revise assumptions. Often, the results of balanced benchmarking will lead to a major rethinking of past assumptions. In the aforementioned study of bank branches, management learned a valuable lesson, that small branches could be among the most efficient operations. That insight helped prevent the bank from making a huge mistake in closing those branches. In another study, a health maintenance organization investigated the efficiency of its 3,000 member physicians with respect to the number of office visits, ambulatory surgery procedures, hospital days, lab and diagnostics tests, emergency room visits and other factors. In order to reduce health-care costs, the HMO was considering cutting the number of its specialists, who were assumed to be less efficient than the general practitioners. But the study found that some specialists who were primary care physicians were providing more efficient care than other general practitioners. This result strongly suggested that, instead of a blanket policy to reduce specialists, the HMO might be better served by a more targeted approach that focused on both specialists and general practitioners who were not using resources as efficiently as their peers.

Don't ignore managerial differences. All other things being equal, a business unit with a manager who is an exemplary leader who inspires his or her staff is likely to perform better than a similar unit with a bad manager. Organizations can use balanced benchmarking to deploy the skills and experience of their best managerial talent to the areas of greatest opportunity. Consider a study of the Department of Supply and Services,[10] a Canadian governmental organization responsible for various purchasing activities. The analysis considered various regional offices of that agency in terms of their cost per contract, volume of contracts per person year, supplies used and so on. At the conclusion of the study, the regional managers were reassigned to improve the organization's overall performance. For example, the person who had been heading a relatively efficient office was transferred to a large regional office that was found to be among the bottom performers. The manager was able to identify inconsistent processing procedures at that office and was successful in decreasing its annual personnel costs by more than $500,000, which was achieved through attrition and transferring staff to other activities within that site.

The Future of Balanced Benchmarking

Balanced benchmarking can be an important component for truly understanding efficiency within any service organization that uses a variety of resources to provide a complex set of services in multiple locations. Service performance may be best evaluated and managed with multiple performance tools, and balanced benchmarking provides invaluable information, particularly when used in conjunction with other measurement systems (such as key performance indicators or the balanced scorecard).[11]

On occasions when we encounter resistance to balanced benchmarking, we often discover that managers of underperforming business units cite explanations to challenge the results of the balanced benchmarking analysis. In some cases, these counterarguments raise valuable points; for example, a management consultant might argue that because no two clients are alike, the input and output measures will have great difficulty in capturing such complexities. However, the purpose of the benchmarking analysis is to take advantage of objective analysis to identify where organizations can improve efficiency. Past assumptions, conventional wisdom, personal experience and relationships to and within the organization are important to consider when managing a business, but these human biases can also cloud management's judgment and a company's potential for improvement. The unbiased clarity brought by balanced benchmarking — an application of "moneyball" to business — identifies critical realities of business we can otherwise easily miss.

H. David Sherman *is a professor of accounting at the D'Amore-McKim School of Business at Northeastern University in Boston.* ***Joe Zhu*** *is a professor of operations and industrial engineering at the School of Business at Worcester Polytechnic Institute in Worcester, Massachusetts. Comment on this article at http://sloanreview.mit.edu/x/54413, or contact the authors at smrfeedback@mit.edu.*

ACKNOWLEDGMENT

H. David Sherman dedicates this article to the late Prof. William W. Cooper, who inspired this work, was a principal creator of data envelopment analysis, and was David's dissertation chair and a dear friend for more than 30 years. Bill Cooper will long be remembered for his creativity, his willingness to advise colleagues and his enthusiasm for developing new scholars, even into his ninth decade.

REFERENCES

1. M. Lewis, "Moneyball: The Art of Winning an Unfair Game" (New York and London: W.W. Norton & Co., 2003).

2. A. Charnes, W.W. Cooper and E. Rhodes, "Measuring the Efficiency of Decision Making Units," European Journal of Operational Research 2, no. 6 (November 1978): 429-444.

3. To apply balanced benchmarking, visit www.deafrontier.net and also refer to J. Zhu, "Quantitative Models for Performance Evaluation and Benchmarking: Data Envelopment Analysis With Spreadsheets," 2nd ed. (Boston: Springer, 2009), chap. 1.

4. D. Grewal, M. Levy, A. Mehrotra and A. Sharma, "Planning Merchandising Decisions to Account for Regional and Product Assortment Differences," Journal of Retailing 75, no. 3 (autumn 1999): 405-424.

5. K.R. Sharma, P.S. Leung, H. Chen and A. Peterson, "Economic Efficiency and Optimum Stocking Densities in Fish Polyculture: An Application of Data Envelopment Analysis (DEA) to Chinese Fish Farms," Aquaculture 180, nos. 3-4 (November 1999): 207-221.

6. G.N. Gregoriou, K. Sedzro and J. Zhu, "Hedge Fund Performance Appraisal Using Data Envelopment Analysis," European Journal of Operational Research 164, no. 2 (July 16, 2005): 555-571.

7. B. Golany, Y. Roll and D. Rybak, "Measuring Efficiency of Power Plants in Israel by Data Envelopment Analysis," IEEE Transactions on Engineering Management 41, no. 3 (August 1994): 291-301.

8. K.W. Einolf, "Is Winning Everything? A Data Envelopment Analysis of Major League Baseball and the National Football League," Journal of Sports Economics 5, no. 2 (May 2004): 127-151.

9. T.R. Sexton, A.M. Leiken, A.H. Nolan, S. Liss, A. Hogan and R.H. Silkman, "Evaluating Managerial Efficiency of Veterans Administration Medical Centers Using Data Envelopment Analysis," Medical Care 27, no. 12 (December 1989): 1175-1188.

10. H.D. Sherman and J. Zhu, "Service Productivity Management: Improving Service Performance Using Data Envelopment Analysis (DEA)" (Boston: Springer, 2006).

11. R.S. Kaplan and D.P. Norton, "The Balanced Scorecard: Translating Strategy Into Action" (Boston: Harvard Business Press, 1996).

i. The efficiency score of 85.7% represents the potential fraction of resources that the inefficient unit can use to become as efficient as a combination of the best practice units. Essentially, the 85.7% rating means the unit can reduce resources by 14.3% [100% – 85.7%].

ii. H.D. Sherman, "Improving the Productivity of Service Business," Sloan Management Review 25, no. 3 (spring 1984): 11-23; Sherman and Zhu, "Service Productivity Management," chap. 2; and Zhu, "Quantitative Models."

Reprint 54413.

MITSloan
Management Review

PDFs ■ Reprints ■ Permission to Copy ■ Back Issues

MIT Sloan Management Review articles, permissions, and back issues can be purchased on our Web site: *sloanreview.mit.edu* or you may order through our Business Service Center (9 a.m.-5 p.m. ET) at the phone numbers listed below. Paper reprints are available in quantities of 250 or more.

Customer Service
MIT Sloan Management Review
238 Main Street E48-570
Cambridge, MA 02142