



Capacity Management in Financial Services

By: Dr. J. Lehr McKenzie, Ph.D.

December 23, 2005

I. Abstract

With the increase in competition, the Financial Services industry is faced with many of the same challenges manufacturing had ten or fifteen years ago. It is no longer possible to view capacity as a problem for only equipment-intensive industries, and Financial Services must gain an understanding of how to manage the capacity efficiency of the human capital employed. The Capacity Efficiency Management Model provides: a vehicle to accomplish this goal; a tool that is compatible with a firm's other performance management tools; and a means for managing capacity efficiency from the strategic to the tactical levels.

II. Introduction

The Financial Services Industry has frequently set itself apart from manufacturing's management methodologies such as activity-based costing, theory of constraints, and capacity management. In general, the industry refuses to consider that manufacturing methodologies would have any application to the services industry. However, as suggested by Bramorski, et al., the time for change may be forced on management by economic conditions.

In recent years, service industries have faced significant competitive pressures resulting from globalization, rapid advances in technology, and high degree of customer sophistication . . . Various modern management philosophies and techniques have been developed and successfully applied in the manufacturing sector. Today's service organizations face competitive pressures similar to those faced by the manufacturing industry during the past decade. Well-publicized business turnarounds in the traditional manufacturing sections, specifically the automotive industry, have encouraged the service organizations to assess the

applicability of solutions successfully applied in the manufacturing sector to address their problems.¹

Although the recent emphasis has typically been on capacity measurement rather than management, capacity management is not new to human capital-intensive economic systems. Metrics are key components of a capacity management system—but they should be viewed as the beginning of the process and not the product. In successful capacity management, the metrics used must align with the overall performance management scheme of the firm so that capacity management becomes one more tool to aid the firm in achieving its strategic goals. Another key component to a successful capacity management plan is a concise understanding of the vocabulary. The definitions used must allow for mutual exclusivity and be consistent throughout time and the firm.

III. Definitions

Organizations have a fixed set of resources dedicated to producing a defined set of outputs (either products or services). Resources can be defined as the people, equipment, and supplies required to produce the necessary output. Capacity may be defined as the ability of an organization to perform work or produce output, and may be expressed in units of time or output. While this definition may serve some useful purpose, it is too imprecise for capacity management purposes and may lead to non-comparable statistics. Therefore, subdivisions of capacity into theoretical, practical, and scheduled provide a better starting point. The Federal Reserve Board (Fed), in measuring capacity as a leading indicator of inflation, works with practical capacity and defines it as, “. . .the greatest level of output each plant in a given industry can maintain within the framework of a realistic work schedule, taking account of normal downtime and assuming sufficient availability of inputs to operate machinery and equipment in place.”² They define capacity utilization in terms of a ratio of actual output to practical capacity.

In microeconomic terms (terms applying to individual business units), the Fed adopts the following definition:

¹ Bramorski, Tom, Manu S. Madan, and Jaideep Motwani, “Application of the Theory of Constraints in Banks,” *Bankers Magazine*, January/February 1997, Vol. 180, No. 1, 53

² Corrado, Carol and Joe Matthey, “Capacity Utilization,” *The Journal of Economic Perspectives*, Winter 1997, Vol. 11, No. 1, Nashville, TN., American Economic Association, 152

. . . capacity output is defined in terms of a full-input point on a production function. Capacity output is the level of output attainable by ‘fully employing’ the variable factors of production, given the current technology and keeping fixed factors at their current levels. . . full-input level of the variable factors means applying such factors without limit, to the point where the marginal productivity of all variable inputs drop to zero.³

This definition of practical capacity is clear enough to avoid any erroneous measurements in capacity management. In simpler terms, variable inputs (such as labor) are added up to the point that by adding one more, input capacity is not increased, and in fact may be decreased. For example, in a service industry such as banking, one can test practical capacity by visualizing a teller line with five windows. The bank manager keeps adding tellers to assist customers. One might assume that two tellers could work efficiently at a window and thus double capacity. At some point, however, as tellers are added, they will get in each other’s way and capacity will begin to decline.

Once practical capacity is defined and agreed upon, it is an easy move to scheduled capacity. This result is guaranteed by the law of diminishing marginal returns. Scheduled capacity is the portion of practical capacity that the firm actually schedules for production. The portion of practical capacity that is not scheduled is labeled marketable idle capacity. The Consortium for Advanced Manufacturing-International (CAM-I) defines idle marketable capacity as capacity where, “a market exists but capacity is idle because of competitor market share, product substitutes, distribution constraints, or price/cost constraints.”⁴ Given these definitions, the basic formula for capacity becomes:

$$\text{Manageable Capacity} = \text{Practical Capacity} = \text{Scheduled} + \text{Idle Marketable}$$

Others argue that total theoretical capacity is the potential productive capacity, given twenty-four hours per day, three hundred sixty-five days per year operation. This is broken down further into other categories. However, unless a firm is actually operating in this manner and these additional hours are shut-down hours for legal or other reasons, it does not add value to spend time analyzing this capacity. It is enough to understand that it exists and is available for long-run

³ Ibid., 161

⁴Klammer, Thomas, ed., *Capacity Measurement & Improvement*, Chicago, Irwin, 1996

planning and expansion purposes. Clearly, as one adds shifts and/or increases shift duration, practical capacity begins to approach theoretical capacity.

Capacity Management

Moving from capacity measurement to capacity management, more definitions become necessary. Scheduled capacity is broken down into *productive* and *non-productive*, both of which are further divided (these definitions will be covered within the capacity measurement discussion).

Capacity as a management metric has been used, at least implicitly, since the shift from a self-sustaining system to a barter system, (as defined in the History of Capacity below) and perhaps was the impetus to shift to a barter system.

It has now come full circle. Initially, capacity addressed how to increase *human* capital productivity and efficiency, and later, how to develop *machines* with greater production capability. The late 1950s and early 1960s saw the onset of the computer era, and with it, the continuous need to increase memory capacity. In the 1980s and 1990s, as the nation moves toward a service-based economy, management is again faced with understanding and improving human capital capacity, or efficiency. As the capacity cycle progresses, various methodologies are used to quantify capacity. Capacity is measured in terms of units, time, and costs. However, the questions facing management today are:

- In view of the modern industry lean operations, is measurement of capacity sufficient?
- In a service industry, human capital is acquired for intellectual power more than muscle power, so how do we deal with the typical capacity measurements in this environment?
- How does capacity fit into my performance management model?
- How does my excess capacity affect my product pricing?

This chapter will provide implicit and explicit answers to these questions as well as a synopsis of a service industry capacity model.

IV. History of Capacity

A. Agriculture

While economic cultures were still developing, human beings were self-sustaining. In a short time however, they learned that in order to sustain the family, they needed to produce only their most efficient goods while trading for other needs, thereby increasing capacity. Later, people learned that products could be moved and traded more efficiently by water and capacity again increased. At this basic level, the improvements in quantity of capacity came through improvements in efficiency. It was clearly more efficient for a person to load a barge with grain and send it downstream for trade than it was to carry it container by container to the barter location.

B. Industrial Based

A view of history shows this pattern continuing as the United States moved from an agricultural economy, to an industrial economy, to the current service economy. The law of diminishing marginal returns allowed the farmer to add only so many people to farm an acre of land before they began getting in the way of one another, decreasing the yield. In agriculture, common sense dictated that to increase the capacity per acre, add machinery to make each worker more efficient. Therefore, machinery was a necessary means to increase efficiency and capacity per acre. Eli Whitney discovered the cotton gin to be a more efficient method of processing cotton than human labor. Adam Smith, in *The Wealth of Nations*, found the specialization of labor in the pin factory to be a more efficient use of capacity. In producing a certain part of a pin and then assembling the parts, the overall factory capacity was increased as each unit of labor became more efficient.

As the economy moved towards industrialization, the focus of capacity moved from how to make human capital more efficient to how to make equipment more efficient. Management focused on reducing downtime, waste, rework, and set-up time. The worker was taken more for granted. It was assumed if the machine was functioning properly, the worker would make it produce the advertised number of widgets. Soon, the goal became to develop as lean a plant as possible while maintaining the production numbers and quality. This created a focus on equipment capital versus human capital, leading to machines with more internal capabilities, requiring less human labor.

Human Capacity Measurements

Clearly, in the service industries, the main factor of production is human capital. The current environment of a service economy dictates that the management of capacity move again to focus on the efficiency of human capital. Unlike human capital use during the periods of rapid economic development, today's human capital is primarily tasked with thinking, and is paid for thought output. Given that a firm purchases the most capable thinker in the industry within its budget constraints, an increase in capacity is directly correlated with an increase in efficiency. A firm may still be able add capacity by purchasing more factors of production, but this is limited by at least two factors that are fixed in the short-run: plant size and money capital. The competitive leaders are those who maximize the efficiency of their factors of production to get maximum goods output from existing capacity.

In order to maximize efficiency of human capital in the services industry, it is useful to think of the employee the same way machinery is thought of in the manufacturing environment. An advantage of human capital over equipment is that if managed properly, human capital does not have to be a depreciating asset. If not managed properly, however, human capital will become obsolete over time. The firm must think in terms of providing inputs to the employee while expecting certain outputs. As management views the ratio of input to return, these inputs become part of the measure of efficiency and management of capacity. Rather than material inputs, (although some material inputs may be required in the form of pens, pencils, computers, etc.), the inputs are mostly going to be intellectual, (such as training, counseling, operating instructions, etc.). Measurement of the efficient use of these inputs is necessary for capacity management. The question of return on investment is raised here just as it is for the overall firm.

The reason for focusing on human capital efficiency measurement versus equipment efficiency measurement in the financial services industry is clear. In a technologically advanced firm (such as most financial services firms), the equipment used by employees typically has capacity capabilities greater than the ability of the user to consume the capacity. For example, most automatic dialers used for collection purposes can dial more calls than the collector can handle in an eight-hour period, and most computer programs run at speeds much higher than the input capacity of the user.

V. Capacity Measurement

Traditional

Traditional measures of capacity have included theoretical, practical, and scheduled. Most practitioners have not worried with theoretical or even practical; instead, scheduled capacity has been the emphasis for planning and budgeting. When measurements are accumulated and used, the common practice has been to measure practical capacity and actual capacity utilized, and then determine whether or not excess capacities exist. The methodologies for collecting measurements of capacity have varied. In many cases, surveys have been used and accepted as a fairly accurate methodology. The Fed has used this methodology since the 1960s. Carol Corrado, of the Board of Governors of the Federal Reserve System, and Joe Matthey, Senior Economist, Federal Reserve Bank of San Francisco, state:

Economists may debate how capacity should be defined, but those who discuss production capability with plant managers quickly discover that managers generally are quite precise about how much their facilities can produce without extraordinary efforts. The Federal Reserve's capacity statistics are grounded in survey evidence on utilization collected at the plant level for the fourth quarter of each year (for the most part).⁵

Other methodologies include both manual and electronic record-keeping by employees and management. Relying completely on employee record-keeping becomes problematic in measuring efficiency, however, as employees are often reluctant to document some needed information (such as time spent on personal business or phone calls). Regardless of the methodology utilized to collect data, the scope of measurement for meaningful management of capacity in the financial services industry must be broadened.

Productive and Non-Productive Scheduled Capacity

Scheduled capacity must be subdivided into productive and non-productive. For productive capacity, the CAM-I definition is adopted:

⁵ Corrado, Carol and Joe Matthey, "Capacity Utilization," *The Journal of Economic Perspectives*, Winter 1997, Vol. 11, No. 1, Nashville, TN., American Economic Association, 152

Productive capacity is capacity used to change the product or provide the service. Productive use of capacity provides tangible changes in the product or service that are of value to the customer. . . Productive capacity results in the production of good products. It may also represent the use of capacity for process or product development.⁶

Good products are defined as, “. . . the right product [service], at the right quality, at the right time, for the right customer.”⁷ Productive capacity is not measured directly, but there are rational distinctions made between non-productive, idle, and practical capacity. To manage capacity, measurements of the actions to be managed provide the metrics that indicate successful management policies and actions; actions resulting in the production of good products and/or services do not need management change. This point becomes more clear as the model unfolds.

The breakdown of non-productive capacity as defined by CAM-I is useful, however, the definitions need to be altered slightly to fit the services industries. In the full capacity model, CAM-I breaks non-productive capacity into standby, waste, maintenance, and setups. Subsequently, each of these categories is subdivided for measurement purposes. Standby is measured in terms of process balance, variability, and scrap. Waste is identified as rework and yield loss. Maintenance includes measurements of scheduled, unscheduled, and time.⁸ For measurement purposes, the definitions of these subsets of non-productive must be mutually exclusive and measurable, and defined in the vocabulary of the industry. For these reasons, some of the CAM-I definitions will be altered and/or omitted.

Standby means exactly what it implies. For financial service industries, standby will include process balance and variability. Human capital is idle due to *process balance* or *variability* (while employees are waiting on work). Process balance is a state created when one factor of production has greater capacity than a prior factor of production in the process. This is generally referred to as a bottleneck. For example, in the Item Processing department in a bank, proof operators constantly wait for documents from branches, other departments, etc. While they may have capacity to produce more good product, (i.e. encoded microlines), they are constrained from doing so by a required prior activity in the process. The same may hold true in the

⁶ Klammer, Thomas, ed., *Capacity Measurement & Improvement*, Chicago, Irwin, 1996, 16

⁷ Ibid, 11

⁸ Ibid, 17

insurance industry as underwriters wait to finish processes on a property, or until the loss prevention department completes their survey of the property.

Variability may come from two different sources: customers or vendors. There is customer variability in banking when tellers are waiting for customers to enter their line. However, if the teller is fungible, one must be careful not to measure customer waiting time as non-productive (i.e., the reason the teller is idle is because there are no customers). There should always be a cause and effect link for measurements. In the case of fungible factors of production, management may have to decide the cause and effect linkage. The consumer lending industry provides an example of vendor variability when a loan originator waits for a credit report from a third-party reporting agency.

Waste for the manufacturing industry includes rework and yield loss. While yield loss is possible in the financial services industry, it is considerably more difficult to quantify (and will not be included as a measurement for the efficiency model). For example, how does one measure the production loss of an inept loan officer? In reality, such a situation's impact should be minimal and not bias the model significantly.

Rework, on the other hand, is of concern in services just as it is in manufacturing. Rework amounts to a factor of production producing a good and/or service that does not meet the criteria of a good product or service (as previously defined). In banking, a careless loan officer create this type of loss. If the loan application is not completed properly prior to approval, rework is required and a measurable amount of non-productive capacity is generated. The original work of the loan officer is non-productive—as well as any time spent by the reviewing committee prior to sending it back to the loan officer.

In the manufacturing environment, maintenance may be a key link to non-productive time. Since machines other than computers account for a small portion of capacity problems, services companies tend to disregard maintenance. Computers have come to the point where downtime is minimal and their capacity far exceeds the users. However, for a complete study of capacity, *all* of the equipment capacity issues faced by manufacturing companies must be addressed by financial service firms that have heavy reliance on equipment (computers, phone systems, auto dialers, etc.). Human capital should be approached like other factors of production.

Earlier it was suggested to think of human capital in the financial services environment the same way one thinks of machinery in the manufacturing environment. The area of maintenance is one area where this comparison works exceptionally well. Firms perform both scheduled and unscheduled maintenance on both human capital and capital equipment. Scheduled maintenance

for human capital is defined as scheduled training, and unscheduled maintenance as unscheduled training. Both of these forms of maintenance are measurable.

Scheduled training includes company-sponsored training classes and off-premise training classes such as university work, seminars, and professional workshops; this is the type of training that employees generally schedule well in advance. Unscheduled training includes spontaneous training or instructional times. These may include corrective interviews, on-the-job counseling by supervisors, and non-routine work reviews by a second party. It should be noted that when non-scheduled training is being measured, it is necessary to measure both the provider and the receiver of the training (assuming that the provider of the training is not hired for the sole purpose of non-scheduled training and is ignoring other productive work to accomplish this end). It is difficult to imagine a situation where a person would be productively hired to do nothing but correct others' errors or counsel them on poor performance. If in fact this was a job description, management should seriously consider counting the total output of this individual as non-productive.

Setups, although clearly measurable in the manufacturing industry, may be more difficult to measure in the financial services industry, and should not be overlooked. For example, in the insurance industry, the Industrial Hygiene department may have considerable setup time as they switch from one type of a test to another; this may be particularly true for hazardous chemicals testing. Likewise, there may be substantial setup time required when measuring noise levels at a policy-holder site. The banking and consumer lending industries may encounter setup time when changing the parameters for the auto dialer. Both of these examples

fall into the category of Changeover. Significant increases or decreases in volume may also lead to additional or different setups. The key here is that the effort spent, while necessary to produce good products and/or services, does not directly result in a good product or service.

The above subdivisions of non-productive capacity now yield a new formula for measuring productive capacity. The formula now becomes:

$$\text{Productive Capacity} = \text{Practical Capacity} - \text{Process Balance} - \text{Customer Variability} - \text{Vendor Variability} - \text{Rework} - \text{Scheduled Training} - \text{Unscheduled Training} - \text{Changeover}$$

From the generic measurement perspective, this leaves two areas to discuss:

- What type of measurements should be taken?

- What is the definition of efficient capacity utilization?

There are numerous methods of measuring the various subdivisions of capacity. As previously mentioned, traditional capacity studies measure practical capacity in units of production; deduct actual used capacity, and the difference becomes excess capacity. In many cases, the difference is zero because firms believe they have control over what expected output capacity is, and manage to that level—thus no excess capacity. True or not, this still begs the question of efficiency. For the Capacity Efficiency Management Model (CEMM), the recommended procedure is to measure capacity in terms of time by each of the divisions of non-productive capacity, idle capacity, and practical capacity. This will allow for the assumption that no more capacity can be obtained with the given factors of production operating at the current level of efficiency; provides for a measure of inefficiency; and leads to actionable items for management. Time is easily assigned to activities in an activity-based costing system and time measurements are easily converted to financial and strategic use measurements.

Additionally, measuring capacity in terms of time allows for easy conversion to the language of different users within the firm. For example, the product team wants to know what capacity is to process, securitize, and service loans (i.e., for how many units of each of these does the firm have the capacity?). The accounting team wants to know capacity to pay invoices, close books, and collect funds. The management team wants to understand capacity relative to setting EVA targets, ROI targets, and making capital investments. The human resources team is interested in capacity to hire, administer health care plans, and how to develop incentives to increase capacity for the firm as a whole. The common language for all of these teams is time. Time can be converted to financial and non-financial measures.

There are two parts to the efficiency equation:

- How many factors of production to add to increase capacity?
- How many non-productive activities to eliminate to increase capacity through efficiency?

The efficient level to add factors of production in order to increase capacity is to add factors up to the level at which the marginal cost of adding the additional factor of production equals the marginal benefit derived from adding the factor. The converse is to remove all units of non-productive capacity where the cost of removal is less than the benefit received by the removal, up to the point that the cost of removal equals the benefit of removal. This verifies the fact that

although non-productive capacity is not efficient, the cost of getting to zero defects would be greater than the benefit derived.

The focus should be on developing an efficient operation to meet the needs of users across the firm

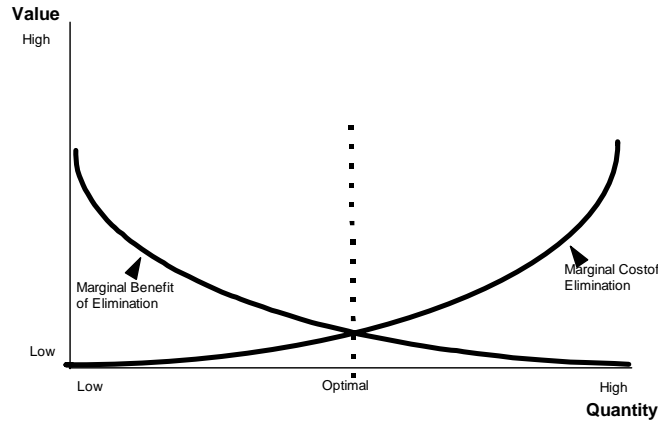


Exhibit 1

With time measurement alone, it is not possible to define the equilibrium point of marginal cost and benefit. However, since employees are either paid by the hour or in a manner easily converted to an hourly rate, it is a simple task to convert the measurements described above into financial measurements. Once financial measurements have been calculated, management has a useful tool for decision-making. The details of what to take measurements against will follow in the discussion of the Capacity Efficiency Model.

Costing Measurement

The majority of developed cost measurements (as with time and unit measurements) deal only with practical and excess capacity, resulting in a dollar value on both. Given this information, management's only opportunities to manage excess capacity are to make decisions that increase demand to use the excess capacity, or to eliminate the excess capacity. This is also limited information for decision-making regarding developing new and/or penetrating existing markets, relative to capacity capabilities. In this type of analysis, there is no indication regarding the efficiency of the scheduled capacity.

For example, an insurance company that decides to expand into three geographic markets may question the capacity capability of the claims department to handle the additional processing. If there is not enough capacity, the company should consider if purchasing capacity is the only alternative or if through efficiency improvements enough additional capacity could be realized to

without hiring additional factors of production. For this answer, the firm needs more information than simply practical and excess capacity quantities. The same scenario may hold true for a bank deciding to add more branches. What is the item processing department's capacity to handle the additional work?

In both cases above, management could make better decisions with a better break-down of capacity. That which appeared to be financially unacceptable and seemed to be the only option (when buying one hundred percent of the additional capacity), may become feasible if a combination of buying and efficiency improvements are possible.

- For example, the data indicates that if the bank adds branches they need 6 more FTE's in the processing department to handle the work flow. Hiring two FTE's would cost \$32,000 each per year. The bank has determined that this would render the project unprofitable. However, the manager of the processing department, having just completed a capacity efficiency study determined they can handle the equivalent of three FTE's work flow by making some changes in their operation. This means the bank only has to fund three additional FTE's. The expansion project is now feasible.

The ability to look at marginal cost of capacity fosters this decision-making process.

Marginal costing

Marginal costing for capacity management in the financial services industry is feasible, but requires careful analysis. The marginal cost to add capacity by investing in human capital is not simply the fully loaded cost of hiring additional employees because there is a difference in the way capacity is consumed in processes, and the manner in which it must be purchased in the market place. A process may only require ten more labor hours, but it may be that labor hours can only be purchased in blocks of forty hours, creating the potential of encountering structural waste. [This topic will be covered in more depth in the section dealing with activity-based budgeting and capacity.] For the present, it is sufficient to understand that capacity cost can be measured marginally and provides management with a clearer picture of the financial implications of various production decisions.

Performance Measures

Having developed unit, time, and cost measures, the next step is to use these metrics to develop performance measures. Capacity measures and management, like activity-based management, kaizen costing, activity-based budgeting, and target costing, are part of the firm's performance management plan. The performance measures should be aligned with performance measures from other areas of management (such as human resources and finance), and become part of the corporate balanced scorecard. Robert S. Kaplan, in articulating the reason businesses need a balanced scorecard, states:

Measurement Matters: "If you can't measure it, you can't manage it." An organization's measurement system strongly affects the behavior of people both inside and outside the organization. If companies are to survive and prosper in information age competition, they must use measurement and management systems derived from their strategies and capabilities. . . The Balanced Scorecard retains financial measurement as a critical summary of managerial and business performance, but it highlights a more general and integrated set of measurements that link current customer, internal process, employee, and system performance to long-term financial success.⁹

VI. Capacity Management

The capacity measurement and management system is a key factor in determining the firm's production capabilities—and should be indicative of the firm's strategies. A correctly constructed capacity model can provide management information regarding capacity use relative to corporate strategy. For example, it is useful for the bank's board of directors to know that sixty-five percent of their human capital capacity is dedicated to assuring high-quality service when their stated vision is to be a low-cost provider of service with average quality.

Kaplan's use of employee measures in the Balance Scorecard takes on particular significance in the financial services industry since the bulk of service production and delivery takes place through the employee. Within the capacity model, methodologies are available for developing individual employee measures and goals as well as team and/or department metrics and goals.

⁹ Kaplan, Robert S. and David P. Norton, *The Balanced Scorecard*, Boston, Harvard Business School Press, 1996, 21

Capacity management is another management tool to aid in strategy achievement. Other tools may relate to continuous improvement, activity-based management, activity-based budgeting, theory of constraints, and others. The capacity tool must be used in conjunction with the other management tools. Used inappropriately or at the wrong time, the results may be the same as a wrench tightening a nut on a screw without a screwdriver to hold the screw in place—effort is exerted but nothing is accomplished except, like the screw, turning in circles.

In addition to having the capability of being an integrated part of the firm's Balanced Scorecard, capacity management supports strategy implementation as well as other efficiency programs the firm may be undertaking (including kaizen costing, activity-based costing, activity-based budgeting, and strategies supported by the survival triplet).

Capacity and kaizen

Kaizen costing is a methodology of continuous improvement undertaken during the production phase of the product or service cycle. Robin Cooper, professor of management at the Peter F. Drucker Management Center at Claremont Graduate School asserts, “kaizen costing systems are most effective when a cost reduction objective is set. . . The aim of a kaizen costing program is to remove unnecessary inefficiencies from production processes.”¹⁰ He also notes, “kaizen costing . . . is deemed inappropriate when the cost of the disruptions caused by the changes to the production process is greater than the savings.”¹¹

Capacity management complements kaizen costing. As previously articulated, a main objective of capacity management is to improve efficiency and result in cost savings. The marginal cost/marginal benefit theory of capacity management mirrors the kaizen costing view of being inappropriate when the cost of the disruptions caused by the changes to the production process is greater than the savings. In assessing the value of capacity management relative to kaizen costing, do not consider a production process as an application related only to manufacturing. A process should be viewed as any series of activities or other processes undertaken in a sequence to produce a predetermined outcome. Using this definition, all firms residing in the service industry perform numerous processes: banks process checks and loan applications; insurance

¹⁰Cooper, Robin, *When Lean Enterprises Collide: Competing Through Confrontation*, Boston, Harvard Business School Press, 1995, 240

¹¹ Ibid, 239

companies process policy applications and claims; and consumer lending institutions process loan applications.

Capacity and ABC

Activity-based costing (ABC) takes the idea of processes a step further. In ABC, processes are identified within, and in support of, the value chain. These processes are then further divided into sub-processes and activities. The purpose of the above is to convert the view of the firm’s operation from that of an organizational view to that of a process view (See Exhibit 2).

Accomplishment of this allows for more accurate measurement of costs, and as will be shown later, a means for measuring and managing capacity.

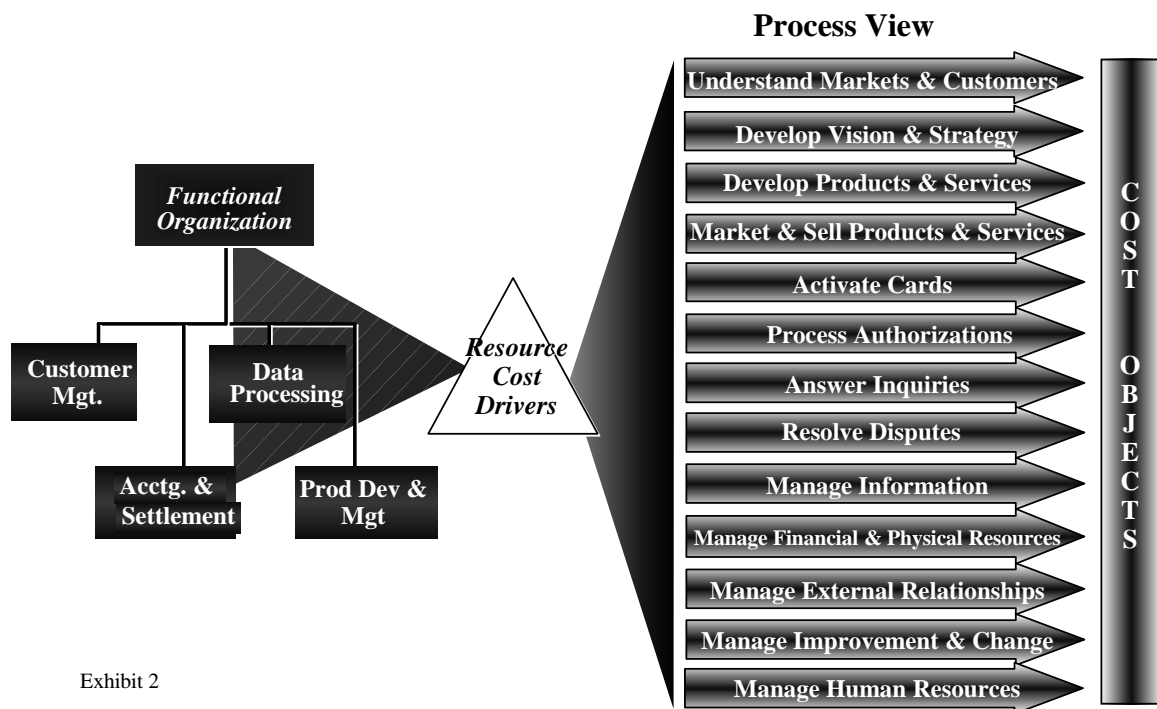


Exhibit 2

Exhibit 2

If ABC is used strictly in the sense of a costing tool, capacity is not a relevant issue for the model itself. Unless there are extreme capacity issues, capacity is just clutter in the system. The

reason for this is the costing procedure is historical in nature, from the **perspective** of developing costs, capacity was what it was **historically**.

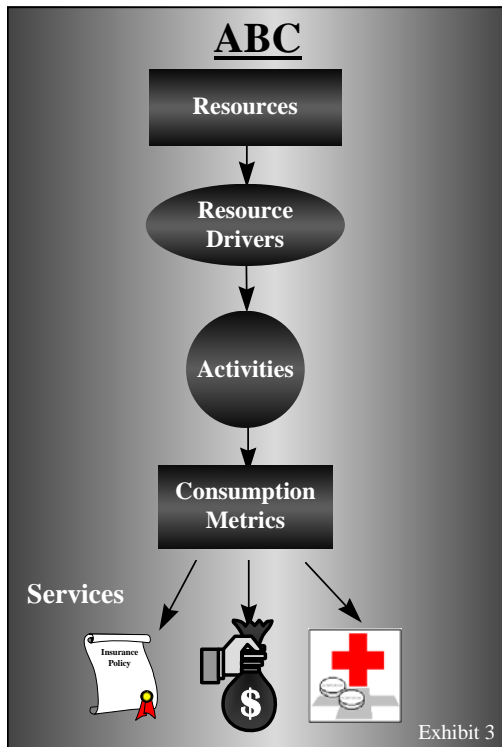


Exhibit 3

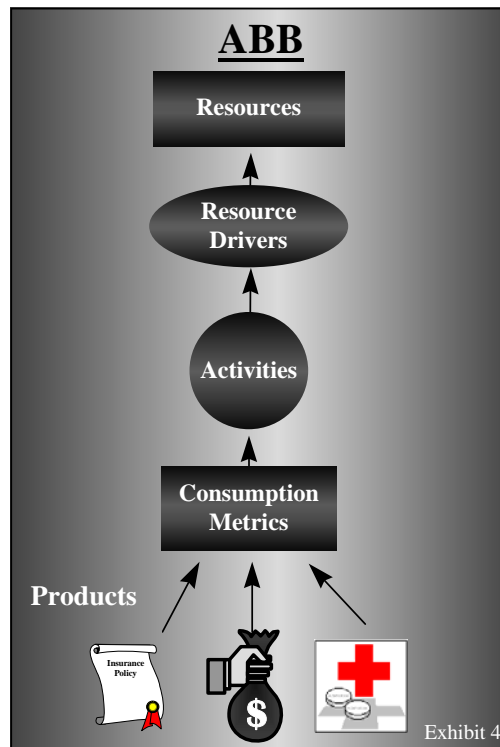


Exhibit 4

While capacity is not a significant tool in the development of an ABC model, an ABC model is a significant tool in the development of capacity models. ABC is a north to south operation that takes resources and drives them down to activities, and then to cost objects (See Exhibit 3). Viewing Exhibit 3 makes it clear that the performance of activities is the driver of product/service delivery. This being the case, capacity at the activity level becomes important. The capacity to deliver products/services is constrained by the capacity of the activity with the lowest output capability.

The activity dictionary provided by an ABC system provides a means for measuring capacity in terms of quantity and efficiency at the activity level. This is accomplished through the ABC model's use of time and financial metrics by activity. Once capacity is measured at this level, it becomes not only measurable, but manageable. This is because managers can manage the activities undertaken by employees of the firm.

Although capacity is not essential to build an ABC model, it becomes important when the goal is to move from activity-based costing to activity-based management (ABM). Management is typically concerned with the historical view, as well as the ability to develop “what-if” scenarios. Once management wants to take this step forward, capacity plays an integral part. For example, if a bank is contemplating a merger with another bank, they must consider the capacity of each of the departments and their respective processes. Is excess capacity going to be created? Will there be a capacity shortage in some areas/processes? Is one bank’s methodology more efficient than the others?

Similarly, if an insurance company is contemplating adding additional agents and/or agencies, capacity questions must be addressed. Does underwriting have the capacity to handle the additional work? Can the claims department sustain the current level of service with the current capacity? Do we need to buy capacity or can we become more efficient and service the additional customer base with the current employee mix? Each of these questions must be answerable by the capacity model. “What-if” analysis leads to questions similar to those encountered in the budgeting process.

Capacity and ABB

Activity-based budgeting (ABB) is a reverse process of ABC (see Exhibit 4), although ABB is dependent on having concise answers to capacity questions. Dr. Robin Cooper believes that there are four things that make ABB complicated; the first is capacity. The problem with capacity in budgeting is it has its roots in structural waste (the difference between the units of purchase for a factor of production and the units consumed in production of the product or service). The reason for this is the fixed and variable nature of factors of production. A factor of production that is variable causes no capacity problems and can be employed on an as-needed basis (e.g. credit reports). Fixed factors of production, from a supply view, create capacity problems if there is a mismatch between the units to be bought and the units it is necessary to consume for production of goods and services (see Exhibit 5).

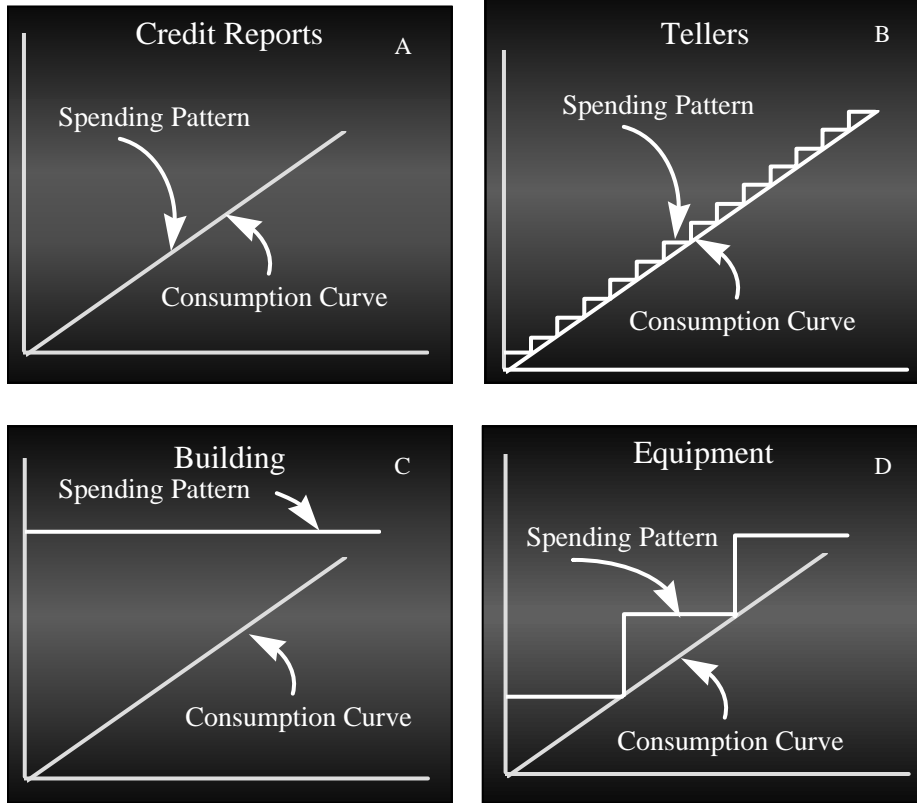


Exhibit 5

Exhibit 5

In discussing these mismatches between supply and consumption, Dr. Cooper stated in an address to the International Management Association:

... , as long as your supply and usage is pretty close together, I can just slide past this little theoretical nicety and not bother to tell you, because you won't spot it. The moment real capacity excesses appear then the underlying knowledge gets crazy. And, in fact, not differentiating between these two [capacity supply and consumption] is the cause of the death spiral because you don't understand capacity. As you go down, [speaking of sales and non-aligned capacity supply/consumption curves] your costs increase. Therefore, you sell less. Therefore, you go down more. Therefore, your costs go up more. Before too long, you're out of business. And, the death spiral is totally caused by a capacity problem. Capacity adjust — you may still go under but, at least, you'll understand

what’s happening to you. So, the cost system won’t save you but, at least, it will warn you.¹²

In Example A, Exhibit 5, there is no difference between the manner in which the factor must be purchased and the manner in which it will be consumed. One credit report can be purchased for each credit application, and one credit report can be consumed per loan application. In Example B, this is not the case. If a bank makes a decision to engage in a marketing effort that is estimated to increase teller usage by 125 customers per day and the current teller capacity is at maximum efficiency, the bank must purchase more capacity. If a teller working at full and efficient capacity can assist 50 customers per day, the bank will need to hire 2.25 tellers. Assuming we cannot hire part-time tellers, this is not possible. The bank will have to hire 3 tellers to create some excess capacity and reduction of profits (see Exhibit 6).

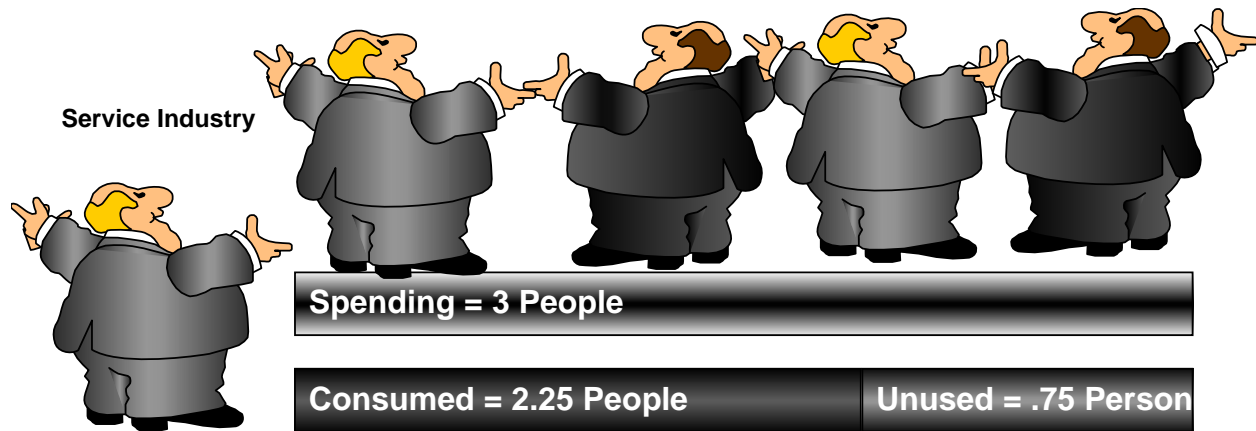


Exhibit 6

Exhibit 6

Example C creates a capacity problem that is not manageable in the short-run. The building size is fixed over some relevant range of time. Without construction of more square footage or purchase of another facility, capacity is fixed. Regardless of what the consumption patterns may be, the supply will remain inelastic. Example D is similar to the human resource example. The difference is that the step variability comes in larger steps than it does for the human resource supply. Computer systems are good examples of this type of supply/consumption variance. An understanding of the various incongruent behaviors between supply and consumption is critical

¹² Cooper, Robin, “ABC/ABM Budgeting Systems: Overview,” Strategic Cost Management Conference II, Institute of Management Accountants, March 12, 1998, Atlanta, GA

in managing capacity. If one does not understand or account for the supply side of the equation, budgeting calculations may be off by considerable amounts.

In budgeting for downsizing, one must be aware of the sticky nature of the supply of production factors. Typically, factors of production are easier to add than to eliminate—particularly in the short run. In the service industries, several examples are applicable. For instance, a collections department is typically staffed to handle a given amount of collection calls in varying amounts based on delinquency status. If there is a sharp downturn in business, capacity is affected in at least two areas: both the automatic dialer and collectors have additional excess capacity. The dialer is very sticky downward. Unless the collections department is going to do away with auto dialing, they have only two options: they can live with the excess, driving unit costs up; or if possible, insource business to use up the excess capacity. The human capital problem is not as difficult, although management is typically hesitant to terminate employees until convinced the downturn is permanent. This leads to excess human capital capacity for a shorter time. Insourcing would be an alternative to use the excess capital as well.

“The Albatross”

The final ABB capacity issue to be discussed is what Dr. Cooper refers to as the “albatross” of unused capacity costs. The Item Processing department in a bank provides a good example. Proof operators have a given capacity to process documents from branches, other departments, etc. If bank management decides to add capacity in anticipation of new branches coming online in two months, the Item Processing department will have temporary capacity. This additional capacity increases the unit cost of processing, but this additional cost cannot be passed on to the existing customers because they will not pay it. Who gets this albatross? Accounting would like to slice it, dice it, and return it to the existing products. Management cannot allow this as it will render the current products non-competitive. The albatross must stay with the originator (marketing, CEO, CFO, or whoever), and until the time the capacity is used, it decreases profitability. The same may hold true in the insurance industry as management decides to add more loss-prevention staff in anticipation of new business brought in by the additional agents hired by marketing.

Capacity and Strategy

Competitive Avoidance and Confrontation Strategy

The tactical use of capacity modeling is not the only application. Capacity can play an informative role in the strategy process as well. One may want to consider capacity as a key

performance indicator in the balanced scorecard for the firm. While strategy can take many different forms in different firms, for discussion purposes, the focus will be on the high-level strategy of confrontation as discussed in *When Lean Enterprises Collide: Competing Through Confrontation*, by Robin Cooper.¹³

The Cooper theory of competitive strategy encompasses several different options. Three strategies are competition avoidance types: collusion, cost leadership, and differentiation. The fourth, confrontation, is developed by competing head-on. In a confrontation strategy, firms do not collude nor do they undertake positions of cost leadership or differentiators. The firm operating in a confrontation strategy will, however, try to develop temporary advantages by lowering price or increasing functionality. The difference between the confrontation firm and the cost leadership or differentiator is that the firm operating in a confrontation strategy does not expect to ever gain a sustainable advantage. Cooper identifies three success factors in confrontation strategy that compose the survival triplet: cost/price, functionality, and quality (Exhibit 7).¹⁴

¹³ Cooper, Robin, *When Lean Enterprises Collide: Competing Through Confrontation*, Boston, Harvard Business School Press, 1995

¹⁴ Cooper, Robin, *When Lean Enterprises Collide: Competing Through Confrontation*, Boston, Harvard Business School Press, 1995, 14

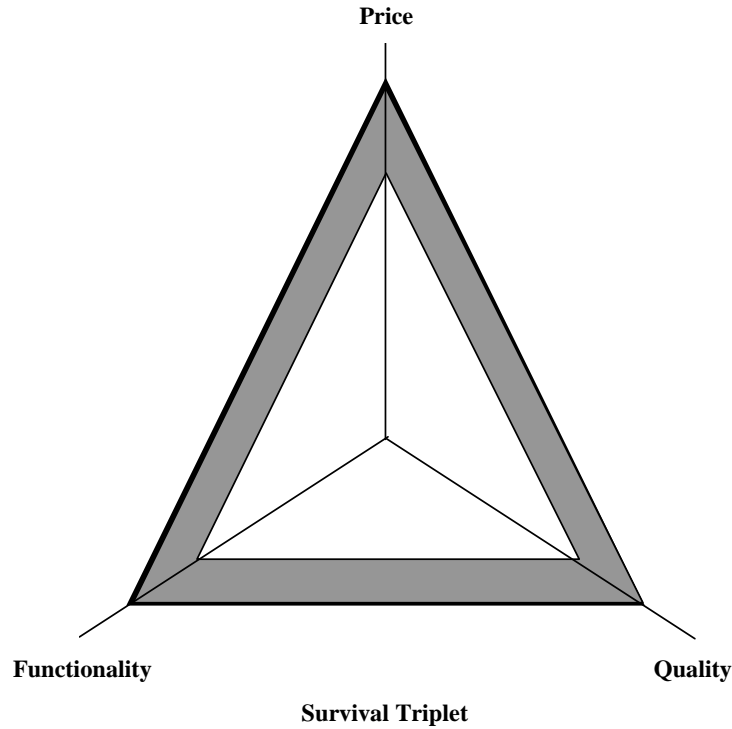


Exhibit 7

Exhibit 7

The survival zone is defined as an area between two concentric triangles created by connecting the maximum and minimum points of each of the success factors, or their survival range. The firm must operate within this survival zone. Under Dr. Cooper’s theory, when the survival zone is large, differentiation and cost leadership strategies are successful; when the zone is small, confrontation strategies are successful. He states, “under confrontation strategy, the challenge is to increase the rate at which the three characteristics of the firm’s products [services] are changing (in the hope of leaving at least one competitor behind) without overshooting customer requirements or becoming unprofitable.”¹⁵

Given the above-stated goal, if one is going to set strategy by manipulating functionality, quality, and price/costs, one must have an understanding of the use of capacity in relation to these factors. Given the assumption that a rational manager will only undertake those activities that are in support of the firm’s stated strategy (or otherwise dictated by law), the question is: what percentage of capacity is assigned to the three factors of success? In an activity-based atmosphere, the answer to this question is obtainable. One simply attributes each activity in the

¹⁵ Ibid, 22

activity dictionary as to its purpose, increase functionality, increase quality, increase price, decrease costs, legal, other.

Once all activities are measured in terms of time per period, the percentage of time allocated from a capacity perspective can be determined. [This will be discussed in more detail later in the chapter.] The result is a management tool that indicates the congruency between capacity use and strategic goals.

VII. Capacity Modeling

Having discussed the historical relevance of capacity management, the various measurement tools, and several management applications, the stage is set to discuss the models available for capacity management in the financial services industry. There are many models that have been used to address capacity issues. The theoretical, practical, and normal capacity-focused models are not necessarily mutually exclusive. In some cases, more than one model has been required to address the specific issues of a firm; other times a single model has addressed multiple issues.

The Resource Effectiveness Model is typically used

in firms using process and assembly-lines in manufacturing, and has its focus in supporting planning and analysis of capacity investments. The Capacity Utilization Model has waste as a focus and divides the causes of waste into time frames and assigns accountability. The design of the Capacity Utilization Model lends itself to use with other continuous improvement-supporting models such as the Theory of Constraints.

Other models designed to accomplish various goals include Capacity Variance Model, CAM-I Capacity Model, CUBE Model, and the Cost Containment Model. Each of these models have elements useful to capacity management in the financial services industry, however, none of them are developed with a focus on services industries. The CAM-I Model is the most comprehensive approach to capacity management and is most conducive to the services industries. The CAM-I model provides a link to the financial reporting system, assigns responsibility for capacity use and misuse, and uses time as a common measure of capacity throughout the firm.

The service firm is more human capital-intensive than the manufacturing firm, and the capacity model needs to reflect such in its measurements. The model developed below, the Capacity Efficiency Management Model (CEMM), is designed to measure and manage capacity relative to

human capital. When capacity management for equipment is necessary, one may use the same model, incorporating the definitions used in the CAM-I Model.

For the CEMM model, the main deviations from CAM-I are the approach to theoretical and idle capacity, and some of the definitions. The CAM-I model begins with the assumption that theoretical capacity is equal to 365 days per year times 24 hours per day. The model then proceeds to divide theoretical into productive capacity, non-productive capacity, and idle capacity. Within idle and non-productive capacity, the divisions become more detailed. While this approach is theoretically correct, it is more detailed for idle capacity than necessary for managing capacity. The focus of capacity management is on the time consumed by production of goods and/or services. All other capacity is potential, and for management purposes it is enough to know it may be available for expansion purposes. Once the strategic decision has been made to expand capacity, one may study the feasibility of using this additional capacity. For this reason, in the CEMM, the only idle time to be addressed is that which occurs during the productive time frame. All other capacity defined as idle under CAM-I will be considered potential capacity (see Exhibits 8 and 9).

The CAM-I Capacity Model

Rated Capacity	Summary Model	Industry Specific Model	Strategy Specific Model	Traditional Model	
Rated Capacity	Idle	Not Marketable	Excess Not Useable	Theoretical	
		Off-limits	management policy		
		Marketable	Contractual		
	Non-productive	Standby	Legal	Idle but useable	Scheduled or Used
		Waste	Process Balance	Variability	
		Setups	Rework	Yield Loss	
		Maintenance	Volume Changeover	Scheduled	
	Productive	Product/Process Development	Product/Process Development	Product/Process Development	Scheduled or Used
		Good Products	Good Products	Good Products	

Exhibit 8¹⁶

Potential capacity as left from the CAM-I Model may be subdivided into that which is constrained by law, and that which is constrained by the firm's management. Examples of legally constrained include all hours that are not available for production due to federal, state, city, county, or other jurisdictions' legal codes. For example, some states prohibit debt collections on Sunday, making twenty-four hours times the number of Sundays in the year the legally constrained hours. The time that is constrained by law is not potential capacity since it cannot be used even if management so desires. Therefore, CEMM addresses only management-constrained potential capacity.

¹⁶ Klammer, Thomas, ed., *Capacity Measurement & Improvement*, Chicago, Irwin, 1996, 17

Firm management-constrained, potential capacity would include all the hours that are available for production but not utilized due to management decisions. These hours would include holidays, shifts not worked (this assumes three eight-hour shifts available to all firms), days off, etc. These are the hours that are not scheduled for production by choice, but are available.

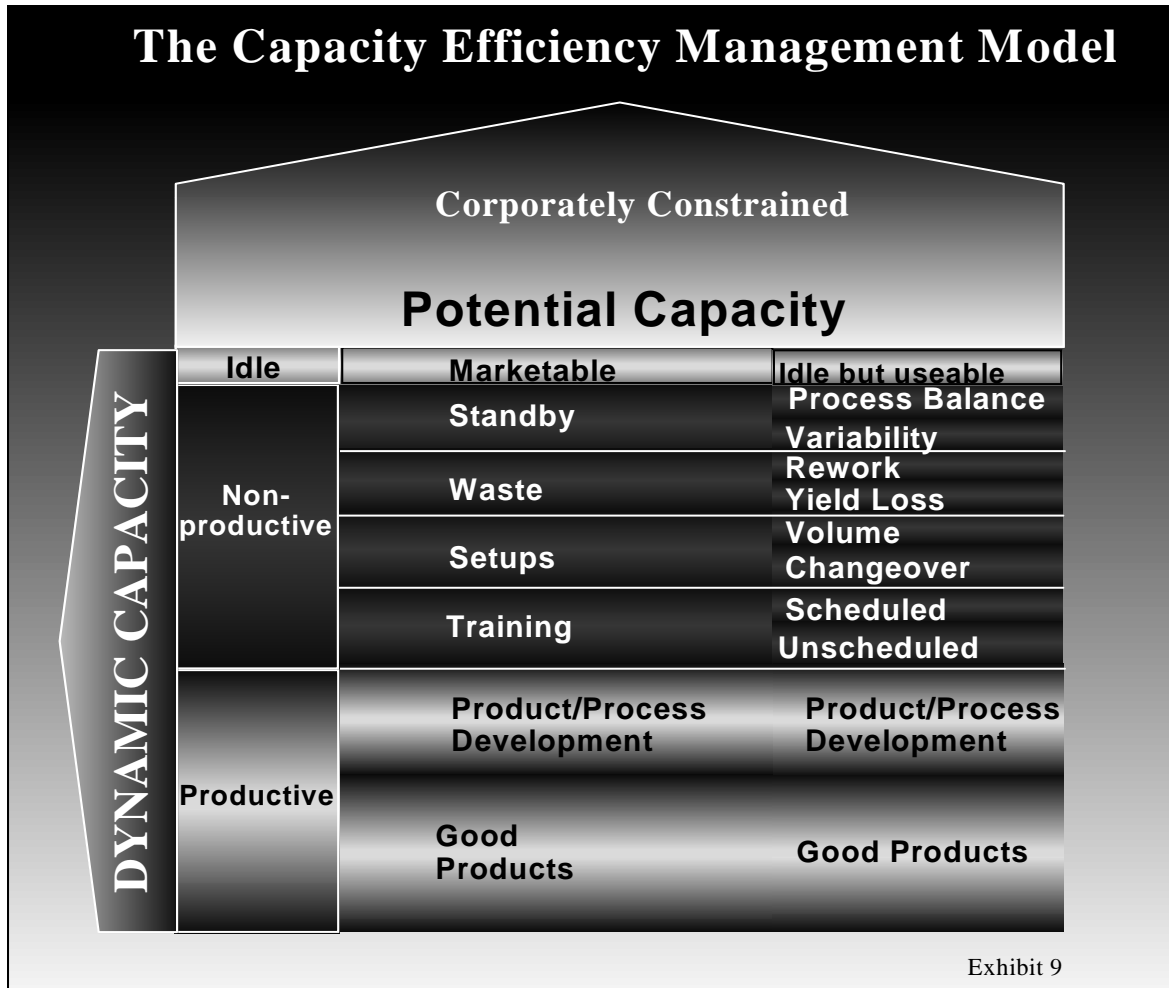


Exhibit 9

The hours that are actively managed are classified as idle marketable, productive, or non-productive. The total of these hours begins with 2080 per full time equivalent (FTE). 2080 hours per FTE is forty hours per week, fifty-two weeks per year. Since most employees do not work forty hours, fifty-two weeks per year, this must be adjusted downward by days off, vacation time, and sick leave, etc. Depending on how one decides to use the model, the calculation may

be accomplished by actual hours per employee, average hours, or weighted averages. This is a management decision based on the need for accuracy versus the cost of accuracy and the availability of data. These hours are captured in the potential capacity area.

Depending on management's needs, potential capacity may be broken down to track the hours considered potential for which the firm is already paying employees (vacation time would be an example of these hours). In theory, a firm may change its vacation policy and move some potential capacity to dynamic capacity. This practice is not uncommon among firms facing union strikes. Some telephone companies have been known to cancel vacations and cross-train management to maintain service capacity during walkouts. Management action moved this capacity from potential capacity to dynamic capacity.

Once dynamic capacity has been defined by subtracting out potential capacity, one may begin the process of measuring dynamic capacity in terms of the more finite divisions. This may be accomplished by several means. One may take actual measurements, as in a time study, or survey methods may be used. If using survey and/or sampling techniques, carefully construct the questionnaires and select participants to avoid internal and external threats to validity. The simplest approach is to first measure idle marketable and non-productive capacity, and then simple subtraction defines productive capacity. Once idle marketable, non-productive, and productive capacity have been measured, productive capacity may be further divided.

To begin the measurement process, activities must first be defined and then measurements made of the various divisions of capacity, activity by activity. Measurement of capacity for CEMM is facilitated greatly by ABC because activities are already defined in a process view—and because one of the natural fallouts of the system is an FTE analysis. Since the initial goal is to measure capacity in terms of time, precalculated Fees provide a vehicle. By knowing how many Fees are involved in any given activity, the human capital capacity assigned to that activity becomes known. Once time per activity per capacity division is calculated, the cost breakdowns may be developed using hourly rates per FTE. These rates may be actual, average, or weighted average—depending on the need for versus the cost of accuracy.

It should be noted that the model totals roll-up moving from right to left; (for instance, scheduled and unscheduled training roll-up to total training and then training rolls-up with set-ups, waste and standby for total non-productive capacity). While all of this seems fairly simple, make sure that double counting does not take place when developing the algorithms.

Once capacity has been measured at the activity level, it may be rolled up to processes, departments, branches, strategy, customers, or whatever cuts management wants to view. The

possibilities are equal to those of the ABC system. While working with an ABC system enhances the capacity modeling, it is possible to model capacity using traditional accounting systems.

In order to use a traditional accounting system, activities would have to be defined within the system. They would probably be defined in a manner consistent with the organizational view of the firm versus the process view. Fees would also have to be calculated for each of the defined activities. Once these two tasks have been accomplished, work can proceed as above. The results of this model will contain the same distortions that one sees when costing using traditional accounting methods.

The measurement process is where many firms stop. They determine if they have excess capacity or if they must buy capacity and move accordingly. While this may be of some strategic assistance, it fall short of the potential of capacity efficiency management. As is true with costing measurements, the value of the capacity measurements comes from the slicing, dicing, and analysis of the numbers. There are numerous possibilities of management reports of capacity. The few reports that follow are from a capacity study for a collections department of a lending institution.

The first management tool to be developed is the overall model reported in various terms such as wages, Fees, man-hours, and units of production.

Selected Model Results: Capacity Allocation for 3-59 Day Accounts

	Manhours	Wages	FTE's	Units
Idle Useable	14,104	\$75,059	32.50	331,159
Process Balance	0	\$0	0.00	0
Variability	22,566	\$120,094	1.04	71,360
Rework	11,283	\$60,047	11.28	35,680
Yield Loss	0	\$0	0.00	0
Volume	0	\$0	0.00	0
Change Over	4,231	\$22,518	14.88	13,380
Scheduled	2,658	\$14,146	0.13	8,405
Unscheduled	113,537	\$604,222	18.19	11,150
Productive	148,961	\$792,740	91.16	73,502
Marketable	14,104	\$75,059	32.50	331,159
Standby	23,272	\$120,094	12.46	73,590
Waste	11,283	\$60,047	11.28	35,680
Setups	4,231	\$22,518	14.88	13,380
Training	116,195	\$618,368	18.32	19,559
Non-productive	154,981	\$821,026	56.93	142,205
Productive	148,961	\$792,740	91.16	73,502
Potential	21,740	\$115,694	110.81	510,444

Exhibit 10

Exhibit 10

The results as formatted in Exhibit 10 yield useful information to management, who can begin to analyze the numbers and ask relevant questions about the operation. In reviewing training, is it efficient to spend only 2,658 man-hours on scheduled training and 113,537 man-hours on unscheduled training? Would a marginal increase in scheduled training result in significant savings in unscheduled training? If the company is on an expansion path, management might want to know if it is possible to reduce both rework and idle useable time by fifty percent, which would free up capacity to make an additional 183,000 to 184,000 more collection calls without adding employees. Other facts raise further questions. Why does the firm's production effort result in more man-hours and wages being dedicated to non-productive rather than productive capacity?

The next view of the data is still at a summary level. This report is one that would be of interest to the CEO, board members, and those who have responsibility for setting the strategic direction of the firm. This model shows that forty-six percent of total man-hours dedicated to the collection process are dedicated to activities with the main function of cost control and/or reduction. If this is the stated firm strategy, there is no problem. However, if the stated strategy is

to be a leader in quality and/or functionality of service, then management must question why the employees are induced towards cost savings activities versus functional or quality activities. This view of capacity is made possible by the attributes assigned to the activities (as previously discussed in the section on confrontation strategies). This model provides a simplified view of the survival triplet. This same view could be presented with wage, unit of production, or Fees as units.

Selected Model Analysis: Strategic Collections' Strategic Use of Man-Hours

◆ 46% of the collection process man-hours are directed towards cost control and/or reduction.

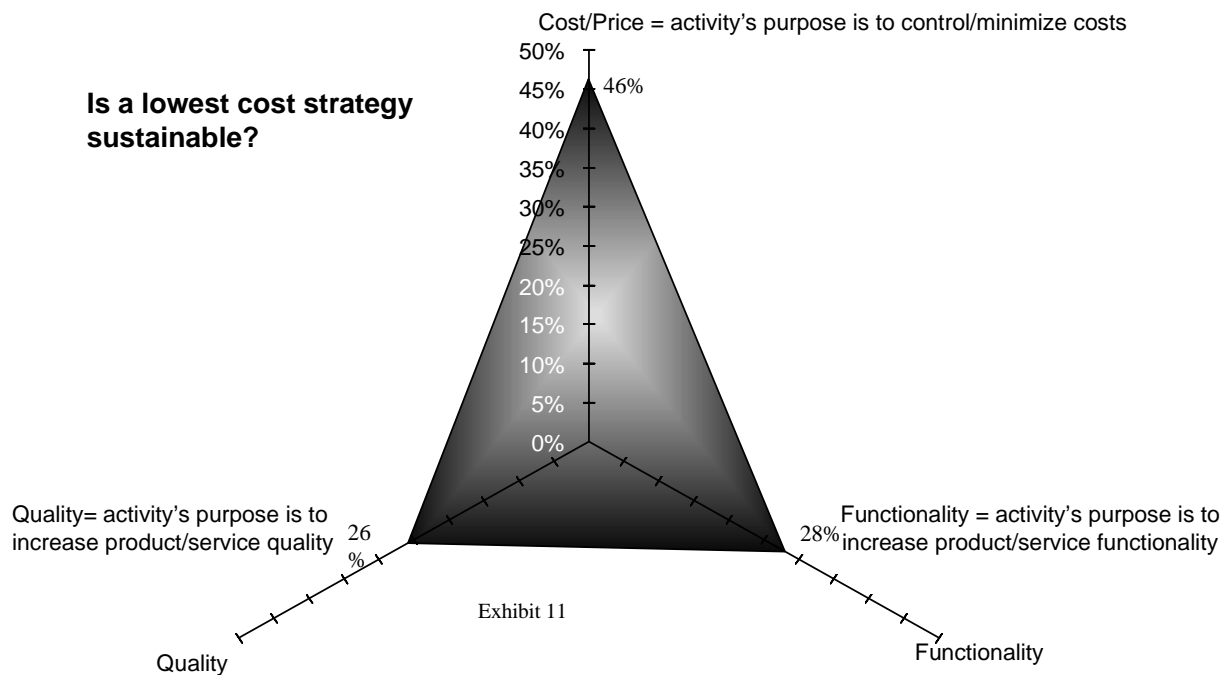


Exhibit 11

Another high-level view (Exhibit 11) gives management a quick assessment regarding what the bottom-line effect might be if certain percentage improvements can be made in converting non-productive capacity to productive capacity. In this model, the improvement rates are set by management, and all other rates are set at the firm's current performance rate. The current number of non-productive attempted calls (calls that resulted in something other than a good product), was taken from the CEMM.

Selected Model Analysis: Converting Capacity to Revenue for 3-59 Day Delinquencies

	3-59 days	3-59 days	3-59 days
Current Number of Idle/Non-productive Attempts	941,000	941,000	941,000
Improvement Rate	20%	35%	50%
Contact Rate Firm	35%	35%	35%
Promise Rate Firm	18%	18%	18%
Kept Rate Firm	45%	45%	45%
Average \$ per Payment Firm	\$325.00	\$325.00	\$325.00
Additional Revenue	\$1,734,028	\$3,034,549	\$4,335,069

Exhibit 12

Exhibit 12

There are numerous other models that may provide information to all levels of management. As a general rule, the higher the management level the report is produced for, the more summarized the report should be. Reports for line management should be very detailed, and would include information from the far right column of the CEMM (this is the level where changes can be implemented at a more granular level). The CEO and Board are more interested in where the firm is directionally (these reports would typically include information to the far left of the CEMM). In addition to reports in a template form as those above, one can also use the data from the CEMM to create effective charts.

The following pie chart (Exhibit 13) provides information at an aggregate level for upper management use (the facts are related in percentages and dollars, and totals are verbalized). In a matter of moments, a CEO or Board can get a feel for the general shape of the collections process.

**Selected Model Analysis: High level management
Collections' Controllable Wages Utilization**

◆ Of \$3,438,906 controllable dollars spent on collections' wages approximately 53% are Idle or Non-productive

**Is there room for reengineering and/or improvement?
Could we enforce new rules which would reduce Idle time?**

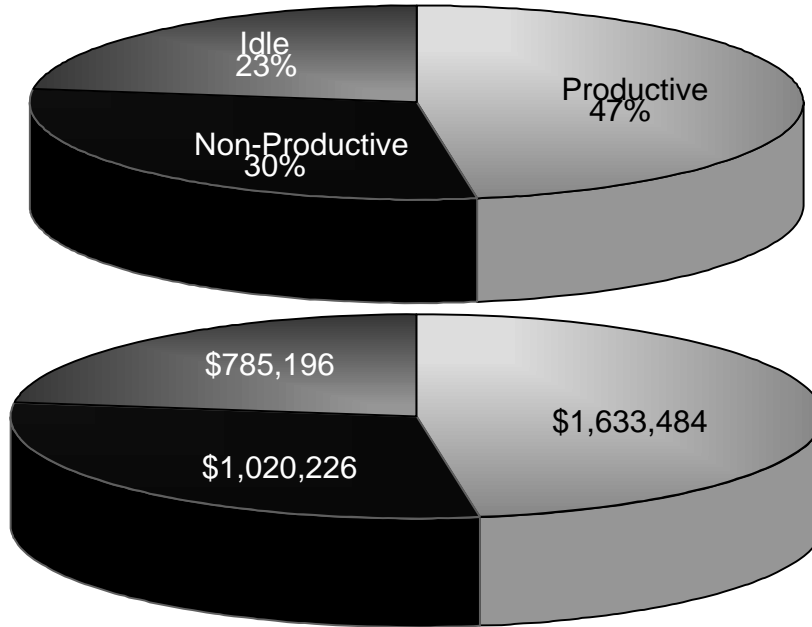


Exhibit 13

Exhibit 13

The next two charts work well together to tell the complete story about counseling capacity use. This was an activity that drew attention on the review of the data in the first general report (Exhibit 9). These charts provide information relating to the capacity allocation for counseling within the collection process. The information is first presented in wages, and then as a percentage of man-hours. Seventy-one percent of the wages spent on counseling in the collections process are spent on those employees with the most experience (the more experienced employees collect the most delinquent accounts). However, only sixty-two percent of the total counseling man-hours in collections is spent in the area with the highest delinquencies. This ratifies the first finding—since the employees working the more delinquent accounts are paid more, one would expect the percentage of man-hours allocated to be lower than the percentage of dollars. This information necessitates evaluating the adequacy of up-front employee training.

**Selected Model Analysis: Improvement Opportunities
Annual Counseling Dollars by Delinquency Classification:**

- ◆ Approximately 27% more dollars are spent counseling employees working on 60+ accounts than on any one of the other three categories.
- ◆ Employees working on accounts with delinquencies between 60 and 120+ days account for 71% of wages spent on counseling in the delinquency activities

Are the additional dollars spent on 60+ employees' counseling necessary and beneficial?

Is up front training and counseling in the 3-59 day category adequate?

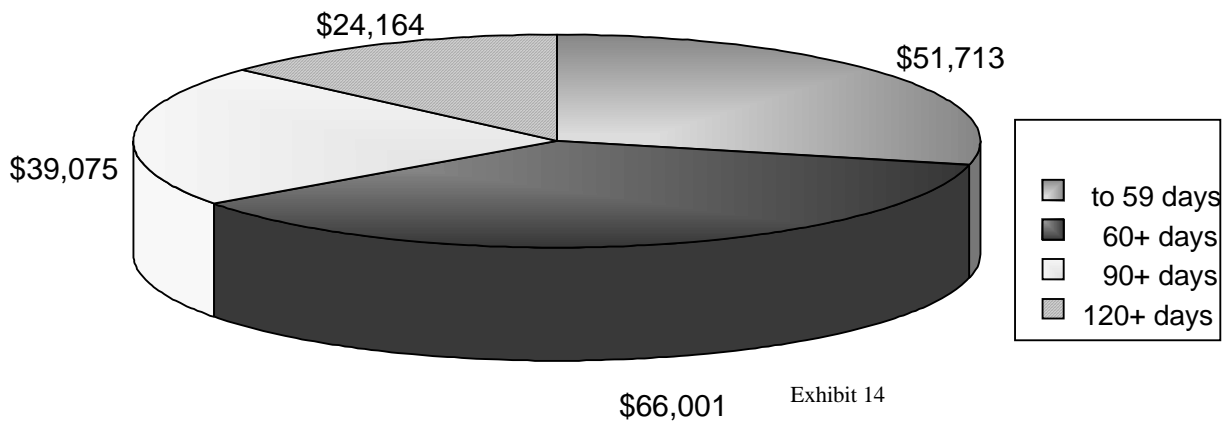


Exhibit 14

Total Counseling Hours % by Delinquency Classification

- ◆ Counseling in the latter stage categories, while representing 71% of costs, represents 62% of man-hours

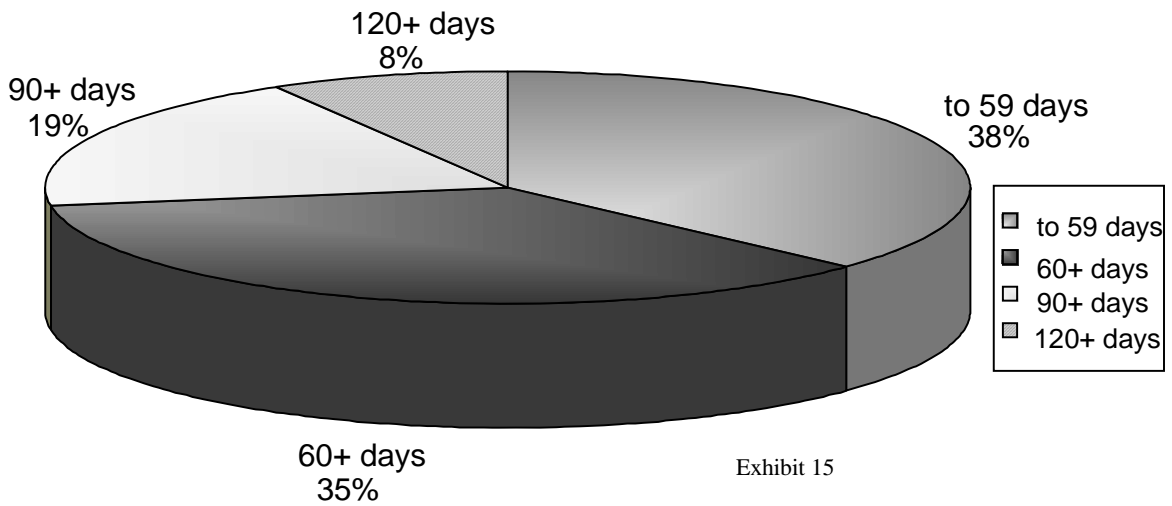


Exhibit 15

As shown in Exhibit 16, more detailed analysis may be shown by pie charts. The format for providing information is open. The data may be presented in pie charts, histograms, line charts, or scatter grams (each type being more or less appropriate for a given analysis).

Selected Model Analysis: Improvement Opportunities
Collection Wages Controllable Detail

◆Of \$3,438,906 annual controllable wages approximately 17% or \$582,306 is spent on idle, but useable time, such as visiting. New rules may apply here and result in benefits

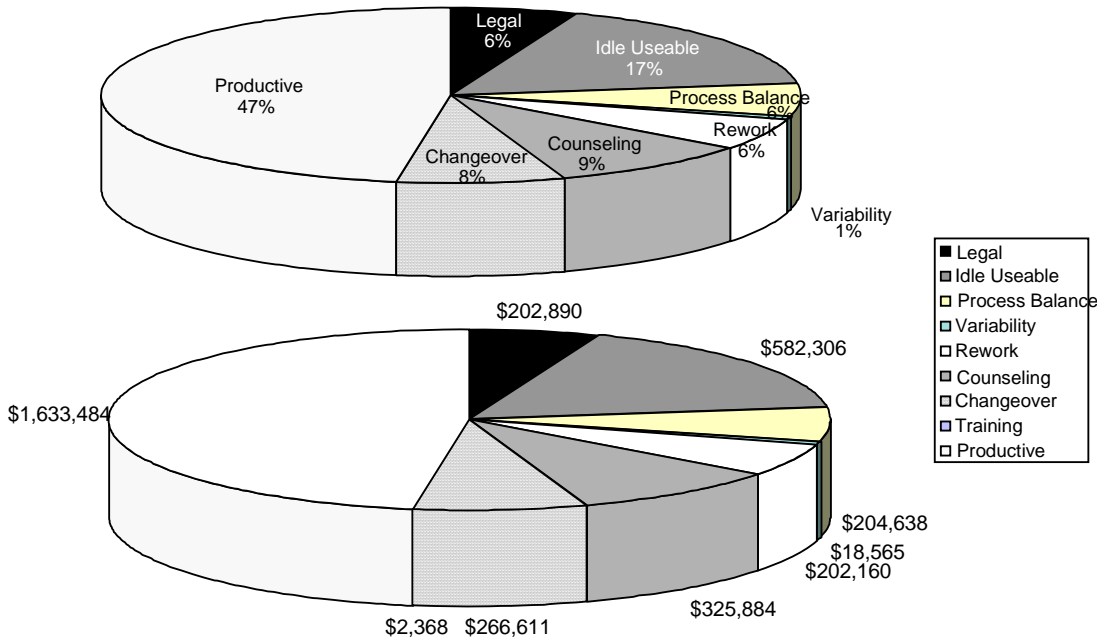


Exhibit 16

Conclusion

Capacity is a management tool that, if measured in detail, may be managed in detail. To simply ask whether or not there is excess capacity may not be enough in a competitive environment. Proper analysis assists management in decisions at the strategic, operational, and tactical levels. When a firm is in a growth mode, it becomes possible to assess the feasibility of capacity efficiency improvement versus capacity purchase. In static times, it can assist management in operating as lean as possible—a clear advantage in confrontation strategies.

Capacity management applies to the services industry as much as it does to the manufacturing industry. It has become just as important for service firms to understand the capacity of the human capital they employ as it is for the automobile manufacturer to understand the capacity of

its robots and other equipment. The capacity management tool should be thought of in the same manner one thinks of the speedometer and fuel gauge of the car. Capacity management is just one more performance measurement instrument to guide the firm along its course.