Firms operating in today’s economic climate are faced with myriad decisions relating to capacity. Few if any studies have considered the long-term effects of firms’ relative capacity positions. In this work, we critically review a broad literature regarding firms’ capacity positions and those positions’ effect on performance over time. We also consider other, less-traditional literatures in their potential to inform the field’s understanding of the capacity-performance relationship.

*I Corresponding Author. E-mail address: cagle_cj@mercer.edu

I. INTRODUCTION

Decisions relating to a firm’s capacity strategy and its capacity position can have important operational and economical effects on long-term performance. A firm operating in a capital intensive industry, for instance, is often required to make capacity decisions involving specialized equipment, large facilities, and technologies resulting in a large commitment of resources relative to capitalization (Porter and Spence, 1978). Further, the time horizon over which these decisions play out likely extends well into the future.

Organizations’ capacity positions will likely vary substantially over time as a result of decisions regarding whether and when to expand or contract. A firm choosing not to expand at a time when others in the industry are doing so may miss important opportunities and greatly compromise its competitive position. A firm choosing to decrease capacity may also discover that it cannot readjust its capacity deficit adequately to meet unanticipated requirements, thereby leading to poor performance. On the other hand, if a firm simultaneously expands with other firms in the industry it may be placing itself as well as its competitors at increased risk to an industry-wide capacity glut (Lieberman, 1987).

A capacity strategy is, in part, used to ensure adequate resources for creative experimentation and innovation (Bourgeois, 1981). Excess capacity, as a strategy, may protect firms by providing buffers for surges in demand (Cyert and March, 1992; Anand and Ward, 2004; Dess and Beard, 1984), or flexibility to seize such opportunities as may emerge (Greenley and Oktemgil, 1998). Capacity strategies intended to bulwark an organization against unanticipated environmental changes may also permit a firm to preempt competitors by exploiting changes in the environment ahead of competitors.

An organization’s capacity position is often the culmination of decision-making processes that involve a number of factors
If an organization carries too much spare capacity, operations may become inefficient. Too little spare capacity may result in a loss of sales or other constraints on action (Hayes and Wheelwright, 1984; Bradley et al., 2010). The specific role that an organization’s capacity cushion takes in the strategic process is often influenced not only by that firm’s particular interests, but also by the environment (Bradley et al., 2010). While some research has been undertaken involving firms’ capacity and their performances at particular junctures (Hendricks and Singhal, 1995; Daniel et al., 2004), research into the ongoing performance consequences of capacity has been limited. Thus, a fuller view of the overall performance effects of firms’ capacity remains largely absent from the literature (Daniel et al., 2004). Table 1 is a summary of more recent empirical studies concerned with the relationship between a firm’s capacity and its performance.

The literature provides multiple justifications for the consequentiality of capacity. For instance, firms may hold excess capacity to allow for a wider range in responses to environmental contingencies (Greenley and Oktemgil, 1998). Further organizations with excess capacity can afford to adopt structures that may not be a good fit for their respective environments because the excess capacity acts to buffer them from the consequences of the mismatch (Cheng and Kesner, 1997). Meyer (1982) also found evidence to suggest that excess capacity may act as a shock absorber for minimizing the effect of reductions in resources.

Capacity plays an important role in firms’ efforts to accommodate demand growth or variability (Olhager et al., 2001). Because capacity changes tend to be lumpy, require large capital investments, and have long lead times, decision makers must carefully weigh the advantages and disadvantages as to when capacity changes are needed (Olhager et al., 2001). According to Hayes and Wheelwright (1984) these changes begin with one of three strategies for responding to anticipated changes in demand: lead, lag, or track (Olhager et al., 2001).

A lead strategy proposes to create excess capacity prior to demand increases. Firms tend to choose this policy if their objective is to never run short (Hayes and Wheelwright, 1984). A second approach, a tracking strategy, attempts to match a firm’s capacity position to its forecasted demand. Its objective is to balance the likelihood of having excess capacity with the likelihood of facing a shortage (Hayes and Wheelwright, 1984). Finally, firms that take a wait-and-see approach often use a lagging strategy. A primary goal for these organizations is to maximize capacity utilization by delaying until absolutely necessary any position changes requiring additional capital investment (Hayes and Wheelwright, 1984).

Although many researchers have depicted excess capacity, a form of organizational slack, as having a positive effect on organizational performance (Cyert and March, 1963; Daniel, 2004), Davis and Stout (1992) and Jensen (1986) found evidence to suggest that excess capacity may reduce or hinder firm performance by burdening the firm with excessive overhead. Yet both Bourgeois (1982) and Greenley and Oktemgil (1998) found evidence to suggest that excess capacity improves organizational performance only to a certain point, beyond which it begins to reduce performance.
TABLE 1. EMPIRICAL STUDIES INVOLVING RELATIONSHIPS BETWEEN CAPACITY POSITION AND FIRM PERFORMANCE

<table>
<thead>
<tr>
<th>Study</th>
<th>Sample/data Source</th>
<th>Capacity Measures</th>
<th>Performance Measures</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short et al. (2006)</td>
<td>2802 firms in 348 industries from 1995 to 2001 (COMPUTSTAT - single firms)</td>
<td>Slack = 7 year average (current ratio)</td>
<td>ROA</td>
<td>A firm's commitment to organizational slack over time was positively related to firm performance. However, slack negatively moderates the rate of performance changes over time. These findings suggest that organizations must balance keeping enough slack on hand to be responsive to opportunities while at the same time limit slack so as not to affect performance growth.</td>
</tr>
<tr>
<td>Cheng and Kesner (1997)</td>
<td>U.S. Airline Industry 1975 - 1979 (Handbook of Airline Statistics by CAB, 1976 and 1979)</td>
<td>Current Ratio = Current Assets/Current Liabilities (Available Slack) Equity-to-Debt Ratio (Potential Slack) Ratio of G&amp;A-to-Sales = G&amp;A/Sales (Recoverable Slack)</td>
<td>Number of Cities Served Number of Planes Ordered Average Coach Fare</td>
<td>Found support for the general prediction that organizational slack has differential effects on a firm's environmental response. Specifically, the relationship between a firm's slack and its response to environmental contingencies is positively affected by a firm's allocation to activities that enhance its operations or marketing effectiveness.</td>
</tr>
<tr>
<td>Hendricks et al. (1995)</td>
<td>128 Manufacturing Firms 1979 - 1990 (Trade and Industry Index (TRND) and Wall Street Journal Index)</td>
<td>Plant Level Capacity Utilization (for year of announcement) Industry Level Capacity Utilization (average 5 years) - (U.S. Census, Survey of Plant)</td>
<td>TABRET (Total Abnormal Return) = day 0 return for firm I, using a market model.</td>
<td>The findings indicated that capacity expansions significantly affect the market value of the firm. However, capacity expansions are negatively related to demand variability.</td>
</tr>
</tbody>
</table>
TABLE 1. - Continued

<table>
<thead>
<tr>
<th>Ritzman and Safizadeh (1999)</th>
<th>144 U.S. Manufacturing firms (32 Job Shop, 46 Batch, 36 Production Line, 30 Continuous) in 15 Industries. Survey participants were selected from the Harris Industrial and Manufacturing Directory.</th>
<th>Capacity Utilization = the reported average annual capacity experienced / theoretically rated capacity</th>
<th>Performance = reported level assessing the relationship between in plant performance relative to other firms in the same industry.</th>
<th>For product-focused plants, the highest performing plants tended to be more capital intensive and have higher capacity utilization. The best performing product-focused plants gain competitive advantage by having increased machine flexibility and larger jobs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Singh (1986)</td>
<td>64 U.S. and Canadian medium and large firms (Moody's Industrial, Transportation, Public Utility and OTC Industrial Manuals; 1974, 1975, 1976)</td>
<td>Absorbed Slack = (Selling + G&amp;A + Working Capital) corrected for transaction volume in sales Unabsorbed Slack = Current Liabilities - (Cash + Marketable Securities)</td>
<td>Net Worth after Tax ROTA after Tax Subjective Performance Index (reported)</td>
<td>Findings reported that good performance is positively related to absorbed and unabsorbed slack. However, unabsorbed slack has no relationship with risk-taking, whereas, absorbed slack is related to increased risk taking. Good companies tended to be associated with higher levels of decentralization. However, firms with high-absorbed slack tended to reduce decentralization.</td>
</tr>
</tbody>
</table>
### TABLE 1. - Continued

| **Greenley and Oktemgil (1998)** | 126 Leading British Firms (63 high performing, 63 low performing) - (Extel Microexstat database) | **Generated Slack:**  
| | | • Cash flow/Investment  
| | | • Debt/Equity  
| | | • EBIT/interest Cover  
| | | • Market/Book Value  
| | | • Current Ratio = Current Assets/Current Liabilities  
| | | • Sales per Employee | **Invested Slack:**  
| | | • Administration Costs/Sales  
| | | • Dividend Payout  
| | | • Sales/Total Assets  
| | | • Working Capital/Sales | **ROE**  
| | | | **ROI**  
| | | | **RONA**  
| | | | **ROS** | **Sales Revenue** | **High Performing Firms:** All measures of generated slack and Sales/Total Assets and Working Capital/Sales from invested slack are significantly related to ROE and RONA. ROI indicated a nonlinear relationship with the Current ratio. ROS was positively related to the Current ratio, Debt/Equity and Sales/Total Assets. Controlling the amount of resources invested in total assets and working capital with respect to sales is important for achieving flexibility that then impacts performance. Low Performing Firms: Neither generated slack nor invested slack are important determinants of performance, suggesting that critical levels of slack required for flexibility are not likely to be found in low performing companies. Overall Differences: High performing firms did not appear to carry more slack than low performing companies. |

| **Bromiley (1991)** | 288 Firms in SIC 3000-3999 (COMPUSTAT and IBES) | **Available Slack** = Current Ratio  
| | | **Recoverable Slack** = SG&A/Sales  
| | | **Potential Slack** = Debt-to-Equity Ratio  
| | | • Interest Coverage Ratio | **ROA**  
| | | | **ROE**  
| | | | **ROS** | **Industry Performance** = average ROA for firms with same 2-digit SIC (code). | **Findings reported that slack appears to reduce risk taking. It is also suggested that slack allows room for risk taking but it is low performance that increases the level of risk taking, not slack per se.**

---

Journal of Supply Chain and Operations Management, Volume 12, Number 2, May 2014
TABLE 1. - *Continued*

<table>
<thead>
<tr>
<th>Study</th>
<th>Sample Description</th>
<th>Measure 1</th>
<th>Measure 2</th>
<th>Measure 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tan (2003)</td>
<td>17,000 Chinese SOE 1995-1996 (archive data)</td>
<td>Absorbed Slack = Capital Depreciation as a percentage of total capital assets</td>
<td>Unabsorbed Slack = Retained Earnings as a percentage of total profit</td>
<td>Both absorbed and unabsorbed slack exhibit a curvilinear relationship with firm performance. It is suggested that there is an optimal level of slack, after which the relationship between slack and performance degrades as the cost of slack increases.</td>
</tr>
</tbody>
</table>
Researchers have long recognized that organizational studies are incomplete without the inclusion of an organization’s environment (Hayes and Wheelwright, 1979; Aldrich, 2008; Porter, 1980; Dess and Beard, 1984; Sharfman and Dean, 1991; Cannon and St. John, 2007). Aldrich (1979) was one of the earliest researchers to suggest that the study of the organization-environment framework is fundamental to understanding organizational structure as well as organizational environments (Aldrich, 2008). Contingency theorists posit that an appropriate organizational form depends on the fit between the organization and its environment (Aldrich, 2008). However, exactly how organizations must vary in response to changes in their environments is often debated (Lawrence and Lorsch, 1967; Aldrich, 2008; Pfeffer and Salancik, 1978).

That environments play central roles in much of the theoretical and empirical research into organizational effectiveness suggests that closer examination of environments would be useful in understanding the performance consequences of actions undertaken by firms (Pfeffer and Salancik, 1978). These actions include the adjustment of capacity.

Organizational environments are by their very nature variable (Aldrich, 2008), with this variation often manifested across multiple dimensions. Tung (1979) noted that the composition of organizational environments is distinctly different from the character of organizational environments, suggesting that organizational environmental studies should incorporate both. As environments become more stable, predictability improves, enabling organizations to develop standardized practices for dealing with environmental elements (Miller, 1988; Aldrich, 2008). Organizations operating in stable environments tend to more accurately forecast demand and assess capacity requirements (Miller, 1988). *Ceteris paribus*, therefore, firms operating in stable environments should tend to outperform those in less stable ones.

As environments become less stable and more turbulent, predictability decreases, affecting the accuracy of demand forecasts and, by extensions, assessments of capacity requirements (Miller, 1988; Aldrich, 2008). Firms operating in these environments often face high levels of uncertainty and unpredictability and therefore tend to be challenged in order to maintain performance over time (Miller, 1988; Olhager et al., 2001).

While environmental complexity has been used to describe the number of interrelationships and elements for planning and control tasks (Van Dierdonck and Miller, 1980), it has also been used as a measure for the number of firms in an industry (Boyd, 1990). Drawing on the Industrial Organization (IO) literature, competitive complexity, a sub-dimension of environmental complexity, is frequently measured as market or industry concentration (Shugart, 1997). When there is high industry concentration or several strong leaders in an industry, competitiveness may remain strong but checked (Porter, 1980). Firms in industries with low concentration can avoid sharing information regarding their competitive position. These conditions tend to encourage firms to preempt competitors or engage in infighting, making the industry less stable (Porter, 1980).

A second sub-dimension of environmental complexity identified by Cannon and St. John (2007) is process/facility complexity, defined differently from earlier definitions of complexity. Specifically, process structural complexity is a composite of three variables, i.e., the levels of mechanization, systemization, and interconnections within and among manufacturing processes (Kotha and Orne, 1989; Cannon and St. John, 2007). Increased process complexity tends to be associated with the more highly automated, capital-intensive processes associated with low cost strategies (Cannon and St. John, 2007). Indications of high process complexity are a high degree of mechanization, systemization, and interconnection. Organizations pursuing an industry-wide cost leadership strategy tend to exhibit high process complexity (Kotha and Orne, 1989).
According to Porter and Spence (1982) the complexities of strategic capacity decisions -- long lead times, lumpy expectations regarding future demand, and the consequences of competitors’ capacity moves -- suggest that the optimal capacity position of individual firms can lead to either significant payouts or staggering losses (Porter and Spence, 1982). Systematically studying the effects of strategy and environmental factors on firms’ performance attempts to take a more complete view regarding capacity and its effects on performance over time. We explore this second facet – the role of the firm’s environment – in the next section.

II. CAPACITY AND THE ENVIRONMENT

Tung (1979) noted that the composition of organizational environments is distinctly different from the character of organizational environments, suggesting that studies involving relationships between organizations and their environments should incorporate both. The composition of environments relates to both internal forces as well as external forces that comprise the focal unit’s environment. Internal forces are those factors that affect the nature of the organization’s products and services such as communication and networking processes within the organization’s boundaries (Tung, 1979). External factors and components, such as customers, suppliers and competitors, make up or compose the external environment (Tung, 1979).

Environmental uncertainty and unpredictability involving external factors are important concepts for organizing and operationalizing much of this research (Kotha and Nair, 1995). Environmental unpredictability is the inability of an organization to accurately forecast the behavior of elements from its environment (Miller, 1988). It is often associated with turbulence or level of disorder and is used in the literature interchangeably with terms such as turbulence, volatility, and uncertainty (Davis et al., 2009). However, turbulence and volatility are more likely to connote a lack of environmental order whereas the term unpredictability is used to refer to the lack of pattern that disorder implies (Davis et al., 2009).

Environmental uncertainty as defined by Pfeffer and Salancik (1978) is the degree to which the future state of the environment cannot be predicted or anticipated. However, uncertainty is indicated by more than change alone. In fact, environments can frequently experience considerable environmental change and still remain predictable (Pfeffer and Salancik, 1978).

Duncan (1972), building on the work of earlier researchers (Emery and Trist, 1965; Terreberry, 1968), suggested that environmental uncertainty could be conceptualized in terms of two dimensions. The first dimension, the simple-complex dimension refers to the degree to which there is a number of factors in a few environmental elements (Duncan, 1972). This dimension captures the essence of Pfeffer and Salancik’s (1978) concentration component affecting the level of environmental uncertainty. Dess and Beard (1984) labeled the simple-complex dimension as complexity.

The second dimension, the static-dynamic dimension, refers to the degree to which the factors in an organization’s environment are stable or changing over time (Duncan, 1972). This dimension is similar to the component affecting environmental uncertainty, referred to as the interconnectedness described by Pfeffer and Salancik (1978). Dess and Beard (1984) labeled the static-dynamic dimension as dynamism.

2.1. Capacity and Environmental Dynamism

Duncan (1972) suggested that organizational environments could be described using two dimensions, the static-dynamic dimension, and the simple-complex dimension. The first of these, the static-dynamic dimension, is the degree to which the organization’s environmental elements remain stable over time (Duncan, 1972). Similarly, Aldrich (1979) suggested that the degree to which an organization’s environment is stable or unstable is
determined by the level of turnover in its environment as well as the extent to which its external environment is being disturbed by a change in the number of environmental interconnections (Emery and Trist, 1965; Terreberry, 1968; Aldrich, 1979, Sirmon et al., 2007). Aldrich (1979) referred to these characteristics as stability-instability and turbulence. Miller (1988) found that as environments become more stable, predictability improves, enabling organizations to develop standardized practices for dealing with environmental elements (Aldrich, 2008).

Dess and Beard (1984) determined that Aldrich’s dimensions of stability-instability and turbulence were similar to Child’s (1972) dimension of environmental variability and could more parsimoniously be represented by a single construct, environmental dynamism. While both sub-dimensions are necessary for conceptualizing and operationalizing dynamism, it is also important to capture the lack of predictability as a function of environmental change (Dess and Beard, 1984; Rasheed and Prescott, 1992).

Although both are facets of environmental uncertainty (Miller, 1980), unpredictability and dynamism are conceptually different in terms of what they measure (Boyd, 1990). Other terms used in the literature to refer to dynamism include turbulence (Tung, 1979), volatility (Bourgeois, 1985), and rate of change (Daft et al., 1988). When decision makers perceive one or more environmental components to be unpredictable, it is often the case that organizations may maintain slack resources – e.g., a capacity cushion – to offset the lack of predictability (Milliken, 1987).

As environments become more stable, predictability increases, enabling organizations to develop standardized practices and routines (Miller, 1988; Aldrich, 2008). Organizations operating in stable environments tend to more accurately forecast demand and assess capacity requirements (Miller, 1988).

As environments become less stable and more turbulent, predictability decreases, affecting not only the accuracy of demand forecasts but capacity requirements as well (Miller, 1988; Aldrich, 2008). Firms operating in these environments often face high levels of uncertainty and unpredictability and therefore tend to require larger capacity cushions (Miller, 1988; Olhager et al., 2001).

### 2.2. Capacity and Environmental Complexity

Environmental complexity is defined by many researchers as the heterogeneity and concentration of an organization’s environmental elements (Child, 1972; Duncan, 1972; Dess and Beard, 1984; Keats and Hitt, 1988). While it has been the focus of many studies including environmental predictability (Duncan, 1972; Tung, 1979), effect uncertainty (Milliken, 1987), and dimensionality (Cannon and St. John, 2007), consensus among researchers as to these considerations remains elusive (Sharfman and Dean, 1991; Boyd et al., 1993, Dess and Beard, 1984; Cannon and St. John, 2007).

Duncan (1972), as one of the first researchers to focus on the relationship between organizations and their environments, identified two dimensions, one of which was used to characterize the perceived level of organizational complexity, i.e., the simple-complex dimension (Duncan, 1972). The simple-complex dimension accounted for similarity/dissimilarity among environmental factors as well as the number of organizational factors (Duncan, 1972). Duncan (1972) also suggested that studies involving organizations’ decision-making capabilities could not be limited to internal components, but should consider external factors such as customers, suppliers, competitors and technologies (Duncan, 1972).

Aldrich (1979), critical of Duncan’s (1972) measures of environmental uncertainty,
suggested two dimensions of the six proposed for characterizing environmental variation (Aldrich, 2008). The first, labeled concentration-dispersion, represented the range over which resources are evenly distributed, not only in organizations’ environments but also in their geographical locations (Aldrich, 2008). The second dimension, homogeneity-heterogeneity, represented the degree of similarity or lack thereof between the environmental elements in the population (Aldrich, 2008). While Aldrich (1979) disagreed with Duncan’s (1972) units of analysis as well as whether to use objective or perceptual measures, Aldrich’s dimensions of similarity-dissimilarity and the number of organizational factors generally correspond with Duncan’s homogeneity-heterogeneity and concentration-dispersion dimensions respectively (Aldrich, 2008).

Building on the work of Child (1972), Duncan (1972), and Aldrich (1979), Dess and Beard (1984) later used factor analysis to determine that Aldrich’s environmental dimensions of homogeneity-heterogeneity and concentration-dispersion formed a more parsimonious construct labeled environmental complexity (Dess and Beard, 1984).

2.2.1. Competitive Complexity

The general consensus among researchers is that environmental complexity involves a measure of the degree of heterogeneity and dispersion of environmental elements (Aldrich, 1979; Dess and Beard, 1984; Keats and Hitt, 1988), and that most researchers have treated environmental complexity as a one-dimensional construct with mixed or inconclusive findings. Complexity, as depicted by Cannon and St. John (2007), is a multidimensional construct consisting of four factors: competitive complexity, market diversity, process/facility complexity, and resource complexity. Of importance is the finding that different measures of complexity capture different dimensions of the relationship between environmental complexity and performance (Cannon and St. John, 2007).

Competitive complexity – competitive diversity – is frequently used as a proxy for environmental complexity (Aldrich, 2008; Dess and Beard, 1984; Cannon and St. John, 2007). As the number and diversity of environmental elements increases, decision makers struggle with their abilities to comprehend the relationships between them (Tung, 1979). Decision makers faced with more complex (heterogeneous) environments perceive the environment to be more uncertain, resulting in a greater need for information and knowledge (Dess and Beard, 1984; Cannon and St. John, 2007). Increased informational requirements may be due to additional organizational activities or more advanced levels of technical or intellectual capabilities needed for understanding between customers, suppliers, technologies, and sociocultural factors (Duncan, 1972; Cannon and St. John, 2007).

Competitive complexity has also been used to describe the number of interrelationships and elements for planning and control tasks (Van Dierdonck and Miller, 1980) as well as a measure of the number of firms in an industry (Van Dierdonck and Miller, 1980; Boyd, 1990). Drawing on the IO literature, competitive complexity is frequently measured as market or industry concentration (Shugart, 1997). According to Stigler (1964), the likelihood of detecting deviations from the norm in an industry is higher when industry concentration is low (Pfeffer and Salancik, 2003). One reason that this may be the case is that industries with a high concentration of firms or that appear to be homogeneous have higher information processing requirements (Keats and Hitt, 1988). Further, when there is high industry concentration or several strong leaders in an industry, such as in an oligopoly, competitiveness may remain strong but checked (Porter, 1980). If there is low industry concentration – i.e., firms of comparable size and resource endowments – one or more organizations may try to preempt competitors, making the industry less stable (Porter, 1980).
2.2.2. Process Complexity

The second measure identified by Cannon and St. John (2007), process/facility complexity, is different from earlier discussions regarding complexity. Specifically, process structural complexity is a composite of three variables, i.e., the levels of mechanization, systemization, and interconnections within and among manufacturing processes (Kotha and Orne, 1989; Cannon and St. John, 2007). Increased process complexity tends to be associated with the more highly automated, capital-intensive processes associated with low cost strategies (Cannon and St. John, 2007).

Organizations that pursue an industry-wide cost leadership strategy will tend to exhibit high process complexity (Kotha and Orne, 1989). Indications of high process complexity are a high degree of mechanization, systemization, and interconnection. Investments in plant equipment and technology are high with a strong emphasis on high volume and low cost. Because redundancy is low, equipment must be more reliable, supply sources more stable and production failure rates lower. These production facilities also tend to have high production standards, low levels of work in process (WIP), and high levels of automation (Kotha and Orne, 1989). Since the objective in this strategy is cost reduction, capacity cushions tend to be minimized (Kotha and Orne, 1989).

Organizations pursuing an industry-wide differentiation strategy are likely to exhibit a low degree of process complexity (Kotha and Orne, 1989). Production processes in these facilities are more likely to involve higher levels of operator control and be subject to higher levels of disruption in material flow. There also tends to be less product standardization and higher levels of raw materials (Kotha and Orne, 1989). Products are more likely to be less standard and, because cost considerations are less of a priority than customer satisfaction, excess capacity is often a consideration.

III. CAPACITY AND PERFORMANCE

Our broad survey of research and practitioner literature suggests that there are two distinct views regarding a firm’s capacity and its performance. The first perspective, capacity position as competitive advantage, is argued to 1) encourage innovation; 2) promote flexibility and responsiveness; 3) reduce cost; and 4) lower risk of obsolescence. Each of these factors, collectively or individually, stands to improve a firm’s performance.

The second view, capacity position as a competitive disadvantage, would seem to suggest that capacity position can result in 1) inefficiencies; 2) satisficing; and 3) lost profits. In this view, a relative capacity position that is misaligned would be expected to erode performance. We explore these important decisions, representing conflicting perspectives more fully in the following subsections.

3.1. Capacity Position as a Strategy

A firm’s capacity strategy, as distinct from its capacity position, involves the pattern of capacity-related decisions or adjustments made over time (Hayes and Wheelwright, 1984). Capacity decisions involve both timing – i.e., when to expand or contract – and volume (Hayes and Wheelwright, 1984). The timing aspect of a capacity adjustment is concerned with balancing forecast demand for capacity with available capacity (Olhager et al., 2001). Using Olhager et al.’s (2001) definition, capacity demand surplus implies that there is insufficient capacity to satisfy demand. A capacity supply surplus, on the other hand, is indicative of excess capacity, such as a capacity cushion.

A capacity strategy is typically classified as one of three “pure” types: lead, lag, or track (Olhager et al., 2001). A lead strategy is based on the idea of a capacity supply surplus in which the objective is to maintain a capacity cushion. For a
lead strategy with increasing demand, capacity is increased in anticipation of increased demand and if the demand trend is decreasing, capacity levels are reduced incrementally to demand levels (Hayes and Wheelwright, 1984; Olhager et al., 2001). A lag strategy is based on the objective of maximizing capacity utilization, i.e., targeting capacity levels to remain at or below demand levels. Capacity is increased as a result of realized demand growth. If the demand trend is negative, a firm’s capacity position is decreased in anticipation of diminishing demand (Olhager et al., 2001). Lead and lag strategies are viewed as pure strategies, and often may not be practical, in which case a tracking strategy is used. A tracking strategy combines aspects of both the lead and lag strategies in order to track demand as closely as possible (Hayes and Wheelwright, 1984; Olhager et al., 2001).

Capacity strategies may be oriented toward competitors, whose moves may be preempted with timely expansion by one or more (perhaps coordinated) firms (Porter, 1980). Capacity cushions serve to protect firms by providing emergency resources as buffers (Cyert and March, 1992; Anand and Ward, 2004; Dess and Beard, 1984), but they also enable firms to seize opportunities by providing for sufficient slack should such opportunities emerge (Greenley and Oktemgil, 1998). Finally, excess capacity can provide firms with the means to develop and conduct critical training capabilities or back-up systems without affecting production operations (Lawson, 2001).

The size of the capacity cushion maintained by an organization is often a strategic decision involving a number of options (Bourgeois, 1981; Porter, 1980). If an organization carries too much spare capacity, operations may become inefficient. Too little spare capacity may result in a loss of sales or other constraint on actions (Hayes and Wheelwright, 1984; Bradley et al., 2010). The specific role that an organization’s capacity cushion takes in the strategic process is often influenced by the environment as well as organizational characteristics (Bradley et al., 2010).

Unfortunately, the literature is not consistent in its depiction of the performance effects of an organization’s capacity position over time.

3.1.1. Capacity Position and Competitive Hindrance

In many industries, investments in capacity expansions are large, costly, and irreversible (Porter, 1980). Capacity typically must be added in large increments, often raising the likelihood of overcapacity and price-cutting (Porter, 1980). For many firms, therefore, capacity cushions are difficult to justify in the short term (Cyert and March, 1992). Organizations may view capacity cushions as wastes of resources rather than as buffers against uncertainty (Nohria and Gulati, 1996). Further, in environments marked by resource scarcity, organizations are likely to be penalized more heavily for having more capacity than necessary (Hayes and Wheelwright, 1984). Decision makers who find themselves with capacity cushions may be more inclined to satisfice rather than optimize (Cheng et al., 1997), and they may structure their firms to further personal agendas rather than wealth creation (Child, 1972; Cheng et al., 1997). Decision makers, acting as agents of the firms’ shareholders, will often act to maximize their own personal interests at the expense of the shareholder’s objectives (Jensen and Meckling, 1976).

Nohria and Gulati (1996) suggested that while too little slack discourages innovation by overly constraining experimental projects, too much slack could encourage a number of other issues. Cyert and March (1963) suggested that capacity cushions might encourage managers to relax discipline involving project approval for such projects as pet projects or other initiatives of questionable value in spite of increased risk or marginal returns. Alternatively, capacity cushions may encourage reduced penalties for early termination or poorly executed projects as a result
of lost interest or boredom (Nohria and Gulati, 1996).

3.1.2. Capacity Position and Competitive Advantage

A larger capacity cushion may translate into competitive strength by enabling an organization to be more responsive to customer demand (Hendricks and Singhal, 1995). Research has suggested that a larger cushion may facilitate research and development (R&D) that can result in new products, services, processes and procedures often leading to valuable opportunities (Cyert and March, 1963; Bradley et al., 2010; Nohria and Gulati, 1996). Cyert and March (1963) found that firms making significant technological improvements tended to carry large cushions funded by earlier successes.

Lead-time is often reduced by increasing capacity (Hendricks and Singhal, 1995). Disruptions caused by outages, unexpected demand surges, and rescheduling can often be mitigated by capacity cushions (Hayes and Wheelwright, 1984). The prospect of quicker or more consistent delivery without the use of overtime or disruption is considered reasonable justification for the expense of the additional capacity cushion.

A capacity cushion as a buffer insulates the organization from external as well as internal variation (Cyert and March, 1963; Sharfman et al., 1988). Cushions have also been shown to reduce organizational infighting or other types of internal political behaviors by providing more resources (Cheng et al., 1997; Bradley et al., 2010).

Hayes and Wheelwright (Hayes and Wheelwright, 1984) suggested that the risk of having a capacity cushion in a growing market is quite low, as excess capacity may be quickly deployed. Additionally, a capacity cushion may enable an organization to attract more opportunity in a growing market especially if competitors are resource constrained (Hayes and Wheelwright, 1984). Hendricks and Singhal (Hendricks and Singhal, 1995) determined empirically that capital markets favorably view firms announcing capacity expansions.

The preceding sections suggest that the research literature is of two minds with regards to capacity position and performance. On the one hand, a capacity cushion has been seen as an important competitive weapon. On the other hand, overly large cushions are seen as deleterious to performance. In much of the research literature, however, the dynamics of capacity and performance are unexplored even though capacity is almost universally treated as an asset whose value is realized over time.

Our explorations here lead us to propose at least a partial reconciliation of these views by considering the distinction between financial and non-financial performance. We do not make the argument that these are separate facets of performance vis-à-vis capacity so much as we suggest that there is a temporal element that must be considered. As we note, capacity cushions are depicted as being especially consequential with regard to innovation or customer responsiveness (Cyert and March, 1963, Hendricks and Singhal, 1995, Nohria and Gulati, 1996, Bradley et al., 2010), yet these benefits are largely depicted as playing out over time. The capacity to pursue and exploit such opportunities, however, must be paid for (in the form of capital/overhead costs) in the present (Cyert and March, 1992, Nohria and Gulati, 1996). Indeed, the bill for a capacity cushion may come due long before the benefits of that cushion fully materialize.

Alternatively, not having enough of a capacity cushion might lead to costs that manifest themselves only in the long run. The costs of an inadequate capacity position – in foregone opportunities, lost goodwill, overly utilized productive assets – likely are difficult to link directly with a particular set of cash flows. Further, from a short-run financial perspective, not having enough capacity could appear to be attractive, since the firm’s overhead burden is relatively lower than it might otherwise need to be.

Our review of the literature suggests that this reality – that decision-makers must carefully
weigh both short-term and long-term benefits, taking care not to allow the more concrete financial consequences (which predominate in the short-run) to obscure very real but not-yet-quantified benefits – is at best only touched upon by the capacity-performance literature. The work of Hendricks and Singhal (1995) is an example of research that offers some guidance in this question, but their work focused on capacity expansions’ “signaling” effect on firms’ market valuations.

IV. CAPACITY POSITION AND PERFORMANCE

As we noted in the previous section, because of sizing, location, and investment issues, capacity adjustments often involve large changes either well before or well after demand actually changes. The timeframe for a firm’s capacity position, therefore, generally is substantially longer than simply that period of time over which capacity is being adjusted. It likely is the case, then, that multiple periods of organizational performance would typically be necessary to assess the long-term effects of capacity decisions.

Generally, both practitioners and scholars alike agree that most organizations operate with at least some slack such as a capacity cushion (Cyert and March, 1963). Evidence in favor of a capacity cushion acting to boost performance comes from Greenley and Oktemgil (1998), who found that high-performing British companies tended to carry large capacity cushions but lower-performing firms did not. In an earlier study, Singh (1986) determined that the indirect relationship between good performance and risk taking was mediated by organizational slack, which, he argued, allowed for greater risk-taking. In that vein, Greenley and Oktemgil (1998) focused on the minimal level of slack necessary to maintain flexibility and found that, for low-performing firms, carrying less than that minimum was not a determinant of overall performance.

FIGURE 1. CAPACITY STRATEGIES (a. Lead; b. Lag; c. Track)
The literature suggests that one reason high-performers hold more slack than low-performers is that slack allows for a wider range in responses to environmental contingencies (Greenley and Oktemgil, 1998). Further, organizations with large capacity cushions can afford to adopt structures that may not be good fits for their respective environments because the excess capacity acts to buffer them from the consequences of the mismatch (Cheng and Kesner, 1997).

Capacity cushions play an important role in firms’ efforts to accommodate demand growth or variability (Olhager et al., 2001). Because capacity additions or reductions typically are not incremental, decision makers weigh the advantages and disadvantages as to when capacity changes are needed, i.e., whether prior to or after demand emerges (Olhager et al., 2001). According to Hayes and Wheelwright (1984) there are three strategies for responding to anticipated changes in demand: lead, lag, or track (Olhager et al., 2001). A lead strategy creates capacity ahead of anticipated demand, lag anticipates decreased demand, and track anticipates no change in demand but attempts to match capacity levels with demand as closely as possible.

A leading strategy proposes to create excess capacity in anticipation of increased expected demand. As depicted in Figure 1a, a capacity addition, C, is incorporated to act as a buffer or cushion in period n prior to demand, D, being realized in tn+k. One reason that a firm may be motivated to create a capacity cushion prior to demand is to reduce the likelihood of outages in the event of unexpected surges in demand or unexpected disruptions in supply (Hayes and Wheelwright, 1984). Additionally, adding capacity in anticipation of demand in a growing market may enable a firm to increase its market share ahead of its competitors who may be more concerned with short-term profitability (Hayes and Wheelwright, 1984).

Firms that take a wait-and-see approach often use a lagging strategy. A primary objective for a firm operating under this scenario is to maximize the average output rate given its design capacity by minimizing its capital investment in excess capacity (Hayes and Wheelwright, 1984). Delaying increases in its capacity position until demand materializes, as shown in Figure 1b, often provides a higher average rate of return on manufacturing investments (Hayes and Wheelwright, 1984). Since excess capacity is likely to be utilized almost immediately in a growth market (Hayes and Wheelwright, 1984), the benefit of a buffering effect from variations in expected demand or other disruptive events may be reduced. Additionally, waiting for demand to be realized before making adjustments to a firm’s capacity position may be viewed as a more conservative approach that could lead to a loss in market position (Hayes and Wheelwright, 1984).

Figure 1c, a tracking strategy, depicts the scenario in which a firm attempts to match its capacity position with the average expected demand over time (Hayes and Wheelwright, 1984). The objective for a tracking policy is to balance the likelihood of having excess capacity with the likelihood of facing a shortage (Hayes and Wheelwright, 1984). In a growth market, a firm would attempt to anticipate long-term capacity requirements and make adjustments accordingly. If, however, the demand does not materialize, or is realized in a different period, the firm’s ability to recover may be reduced to fewer options (Hayes and Wheelwright, 1984).

4.1. Research Propositions

Slack improves organizational performance but only to a certain point, beyond which it begins to reduce performance. Research has shown that market participants generally view capacity expansions positively (Hendricks and Singhal, 1995), suggesting that the long-term effects of capacity slack are positive. We posit that over time the relationship between slack and firm performance will be positive.
P1: There is a positive relationship between a firm’s relative capacity position and its performance over time.

As we noted earlier, however, this performance effect typically will be realized over time. Further, in the short-run, before the benefits of the capacity cushion will be realized, there will be capacity costs that must be borne. This leads us to break down P1 into two finer-grained propositions:

P1a: In the short run, there will be a negative relationship between a firm’s relative capacity position and its financial performance.

P1b: In the long run, there will be a positive relationship between a firm’s relative capacity position and both its non-financial and financial performance.

As environments become more stable, predictability improves, enabling organizations to develop standardized practices for dealing with the environment (Miller, 1988; Aldrich, 2008). Organizations operating in stable environments tend to more accurately forecast demand and assess capacity requirements (Miller, 1988). Therefore, firms pursuing a cost leadership strategy tend to minimize capacity cushions (Kotha and Orne, 1989) and are more likely to exhibit better performance over time in stable environments.

Conversely, as environments become less stable and more turbulent, predictability decreases, affecting the accuracy of demand forecasts and making it more difficult to make good estimates of required capacity (Miller, 1988; Aldrich, 2008). Firms operating in these environments often face high levels of uncertainty and unpredictability requiring larger capacity cushions in order to maintain flexibility (Miller, 1988; Olhager et al., 2001). Building on these findings suggests that environmental dynamism will positively moderate the relationship between a firm’s capacity cushion and its performance.

P2: Environmental dynamism will positively moderate the relationship between a firm’s relative capacity position and its performance over time.

It should be noted, however, that there are both financial and non-financial facets of performance that should be considered. In considering environmental dynamism, the literature deals largely with the buffering aspect of capacity cushions. Given this, we would anticipate there being more obvious (and therefore more easily observed/quantified) benefits to capacity cushions in dynamic environments. Thus, we propose that, in contrast to our general proposition involving capacity cushions and performance, in the presence of environmental dynamism short-run financial benefits should be obtained.

P3: In the short-run, environmental dynamism will negatively moderate the relationship between a firm’s capacity position and its financial performance.

Complexity has also been used to describe the number of interrelationships and elements for planning and control tasks (Van Dierdonck and Miller, 1980). Additionally, it has been used as a measure of the number of firms in an industry group (Boyd, 1990). Drawing on the IO literature, environmental competitive complexity is frequently measured as market or industry concentration (Shughart, 1997). When there is high industry concentration or several strong leaders in an industry, such as in an oligopoly, competitiveness may remain strong but checked (Porter, 1980). If there is low industry concentration or firms tend to be comparable in size and resources, one or more organizations may try to preempt competitors, making the industry less stable (Porter, 1980).

Therefore, the level of competitive complexity will tend to positively moderate the relationship between a firm’s cushion and its performance over time.
TABLE 2. FACTORS IN THE RELATIONSHIP BETWEEN CAPACITY POSITION AND FIRM PERFORMANCE

<table>
<thead>
<tr>
<th>Factor</th>
<th>(Timeframe)</th>
<th>Performance</th>
<th>Financial</th>
<th>Non-Financial</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capacity Cushion</td>
<td></td>
<td></td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>Environmental Dynamism/Capacity Cushion</td>
<td></td>
<td></td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Competitive Complexity/Capacity Cushion</td>
<td></td>
<td></td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Process Complexity/Capacity Cushion</td>
<td></td>
<td></td>
<td>–</td>
<td></td>
</tr>
</tbody>
</table>

P4: Competitive complexity will positively moderate the relationship between a firm’s relative capacity position and its performance over time.

Organizations pursuing an industry-wide differentiation strategy are likely to exhibit a low degree of process complexity (Kotha and Orne, 1989). Production processes in these facilities are more likely to involve higher levels of operator control and be more severely affected by disruption in material flow. There also tends to be less product standardization and higher levels of raw materials (Kotha and Orne, 1989). Products are more likely to be less standard and because cost considerations are less of a priority than customer satisfaction, excess capacity is often a consideration. We therefore anticipate that process complexity will negatively moderate the relationship between a firm’s cushion and its performance.

P5: Process complexity will negatively moderate the relationship between a firm’s relative capacity position and its performance over time.

Table 2 summarizes our proposals regarding the effects of a firm’s relative capacity position on both its financial and non-financial performance.

V. CONCLUSION

Many researchers view the effect of a capacity cushion on performance as positive (Cyert and March, 1963; Daniel, 2004). Research has shown that market participants generally view capacity expansions positively (Hendricks and Singhal, 1995), suggesting that the long-term effects of capacity cushions are by and large positive. Research has been consistent that facets of the environment in which the capacity is deployed also should be considered. This research therefore echoes this finding and argues that here is a positive relationship between a firm’s relative capacity position and its performance over time.

As environments become more stable, predictability improves, enabling organizations to develop standardized practices for dealing with environmental elements (Miller, 1988; Aldrich, 2008). Organizations operating in stable environments tend to more accurately forecast demand and assess capacity requirements (Miller,
1988). Therefore, firms pursuing a cost leadership strategy tend to minimize capacity cushions (Kotha and Orne, 1989) and are more likely to exhibit better performance over time in stable environments.

As environments become less stable and more turbulent, predictability decreases, affecting accuracy of demand forecasts and capacity requirements (Miller, 1988; Aldrich, 2008). Firms operating in these environments often face high levels of uncertainty and unpredictability requiring larger capacity cushions in order to maintain flexibility (Miller, 1988; Olhager et al., 2001). Building on these findings suggests that environmental dynamism will positively moderate the relationship between a firm’s capacity cushion and its performance.

Complexity has also been used to describe the number of interrelationships and elements for planning and control tasks (Van Dierdonck and Miller, 1980). Additionally, it has been used as a measure of the number of firms in an industry group (Boyd, 1990). Drawing on the IO literature, environmental competitive complexity is frequently measured as market or industry concentration (Shughart, 1997). When there is high industry concentration or several strong leaders in an industry, such as in an oligopoly, competitiveness may remain strong but checked (Porter, 1980). If there is low industry concentration or firms tend to be comparable in size and resources, one or more organizations may try to preempt competitors, making the industry less stable (Porter, 1980). Therefore, the level of competitive complexity will tend to positively moderate the relationship between a firm’s cushion and its performance over time.

Organizations pursuing an industry-wide differentiation strategy are likely to exhibit a low degree of process complexity (Kotha and Orne, 1989). Production processes in these facilities are more likely to involve higher levels of operator control and be more severely affected by disruption in material flow. There also tends to be less product standardization and higher levels of raw materials (Kotha and Orne, 1989). Products are more likely to be less standard and because cost considerations are less of a priority than customer satisfaction, excess capacity is often a consideration. Therefore, it is likely that process complexity will negatively moderate the relationship between a firm’s cushion and its performance.

VI. REFERENCES

Non-fiction.


