Coordination Costs and Firm Boundaries: A Tale of Two Supply Chains in the Apparel Industry

Scott Wallace
University of Wisconsin-Stevens Point

Brian T. Kench
The University of Tampa

Barb Mihm
University of Wisconsin-Stevens Point

The transaction costs literature emphasizes the role of firm boundaries in addressing incentives of parties to engage in rent seeking behavior. The focus on the potential for "hold-up" in the presence of transaction-specific investments implicitly assumes that production and coordination problems have been solved a priori. Drawing from the capabilities literature, this paper looks at the role organization plays in reducing these costs. It argues that many of the hold-up problems associated with "temporal specificity" can be more accurately characterized as issues of suboptimal coordination. A case study examining alternative organizational arrangements in fast fashion is presented in illustrating the importance of firm boundaries in addressing coordination problems.

INTRODUCTION

In their text *Economics, Organization, and Management*, Paul Milgrom and John Roberts assert that economic organizations exist to address problems of coordination and motivation (Milgrom and Roberts, 1992). Despite the recognition of the dual nature of organizations, the economics literature has primarily focused on issues of motivation in explaining the boundaries of the firm. In particular, "theorists focused on problems of contracting under asymmetric information, developing new methods to deal with the strategic use of information and the attendant incentive problems" (Holmstrom, 1999, p.75). Much of this is the result of the work of Oliver Williamson and others during the 1970s and 1980s in expanding on Coase's insights from *The Nature of the Firm* (Coase, 1937; Klein, Crawford, and Alchian, 1978; Williamson, 1985). "Our understanding of firm boundaries has been sharpened by identifying more precisely the nature and sources of transaction costs in different circumstances. In the process, the focus of attention has shifted away from coordination problems originally emphasized by Coase and towards firm boundaries in providing incentives" (Holmstrom and Roberts, 1998, pp.73-74). While Coase stressed the search, negotiation, and contracting costs of market coordination, the modern transaction cost literature has narrowly focused on incentive problems inherent in market relationships when transaction specific investments are called for. Firm boundaries play an important role in providing incentives to make these investments. An alternative incentive-based conception of the firm in the economics literature

that gained prominence in the 1970s focused on agency costs. According to this view, firms enhance economic efficiency by aligning incentives among factors of production (Alchian and Demsetz, 1973; Jensen and Meckling, 1976). Advances within agency theory have more clearly defined the role of firm boundaries in addressing divergent interests (Holmstrom and Milgrom, 1991; Holmstrom and Roberts, 1998).

These approaches have narrowly conceived of firms as organizational structures that solve incentive problems that otherwise would undermine cooperation among resource owners (Langlois and Foss, 1999). Nevertheless, as Milgrom and Roberts have expressed, organizations are responsible for both "motivating individuals so that their goals are aligned (the agency problem) and organizing individuals so that their actions are aligned (the coordination problem)" (Heath and Staudenmayer, 2000, p.155). The emphasis on incentive alignment within the literature has obscured the role of organizations in aligning individuals' actions. This lack of attention regarding coordination has at times led to the misapplication of economic theory and an incomplete view of the economic role of the firm.

This paper supports an alternative approach that focuses on coordination problems involved in production as determinants of the boundaries of the firm. While incentive issues are critical, coordination and production costs have a greater effect on the choice of governance structures. "Seldom if ever have economists of organization considered that knowledge maybe imperfect in the realm of production, and the institutional forms may play the role not (only) of constraining unproductive rent-seeking behavior but (also) of creating the possibilities for productive rent-seeking behavior in the first place" (Langlois and Foss, 1999). Capabilities theory which stresses the knowledge, skills, and experience embodied within firms maintains that it is the transaction costs associated with coordinating diverse, complementary activities that play a dominant role in determining the vertical boundaries of the firm. According to this approach, vertical integration is often a response to the absence of appropriate capabilities on the market to carry out particular activities along a product's supply chain. In this case, firms economize on these transaction costs by undertaking these stages themselves in order to ensure the product's implementation (Richardson, 1971; Langlois, 1992). Incentives obviously matter but only play a complementary role in determining organizational choice. In other words, if organizations are unable to align actions of different resource owners in the first place, individual incentives do not matter much.

In making the case for the capabilities approach, this paper begins by briefly evaluating the transaction cost and agency cost theories that have dominated organizational economics for the last forty years. Next, it provides an overview of capabilities theory, describing the concept of "dynamic transaction costs" and the implications for analyzing the vertical boundaries of the firm. Then, the paper compares and contrasts the notion of "temporal specificity" (a type of asset specificity) with "dynamic transaction costs." It argues that "hold-up problems" associated with "temporal specificity" often can be more accurately characterized as ones stemming from delays arising from suboptimal coordination rather than opportunism. It then goes on to discuss the role of firm boundaries in addressing this coordination problem. In supporting the capabilities approach, the paper, lastly, provides a case study analyzing alternative organizational arrangements in the apparel industry.

TRANSACTION, AGENCY COSTS, AND THE BOUNDARIES OF THE FIRM

As stated above, the dominant approaches in economics focus on the role of organizations in structuring incentives. The modern transaction costs literature asserts that firm boundaries are critical in providing incentives for highly productive, transaction specific investments. The sunk nature of such investments can pose challenges for contractual relationships by creating ex post incentives to renegotiate the terms of trade to capture the resulting quasi-rents (the hold-up problem). If the value of these quasi-rents causes arm's length contracts to be outside the self-enforcing range, vertical integration may be necessary to induce parties to make transaction specific investments by mitigating potential hold-up problems (Klein, Crawford, and Alchian, 1978; Williamson, 1985). Under these circumstances, contractual solutions are impracticable because of the impossibility of anticipating and writing a contingency for every possible future state of nature.

The property rights literature has deepened this analysis by clearly identifying how vertical integration addresses this incomplete contract problem. According to this view, the ownership of physical assets ultimately determines the boundaries of the firm. Asset ownership strengthens incentives for investments in specialized-assets by granting owners 'residual rights of control' thereby eliminating the need to engage in contingent claims contracting. Vertical integration, however, has its costs as well because it simultaneously diminishes incentives for the former asset owners who now are employees of the integrated firm (Grossman and Hart, 1986). The "optimal allocation of property rights – or governance structure – is one that minimizes efficiency losses. Thus, in a situations where party A's investment is more important than party B's investment, it is optimal to allocate property rights over the assets to party A, even if this discourages investment by party B" (Aghion and Holden, 2011, p.183).

The agency cost literature sees firms as addressing incentive problems that undermine the productivity gains associated with cooperative specialization among resource owners. "Two key demands are placed on an economic organization – metering input productivity and metering rewards" (Alchian and Demsetz, 1973). In doing so, "firms help to solve agency problems, whether these emanate from moral hazard (the agent's action or effort is unobservable by the principal) or from adverse selection (the agent's type is unobserved by the principal" (Aghion and Holden, 2011, p.183). The "key characteristic of a firm is the monitoring of inputs rather than outputs. The need for monitoring arises because much output is jointly produced so that individual contributions get disguised by the contributions of other parties. This results in free-riding and a suboptimal level of performance" (Holmstrom, 1999, p.77). Subsequent contributions in agency theory have taken a more expansive approach with firms doing more than addressing shirking in the presence of joint production. Jensen and Meckling argue that contractual "relations are the essence of the firm, not only with employees but with suppliers, customers, creditors, etc. The problem of agency costs and monitoring exists for these contracts are independent of whether there is joint production…." (Jensen and Meckling, 1976, pp.288-89). The firm essentially represents a nexus that "serve[s] as a focus for a complex process in which the conflicting objectives of individuals (some of whom may 'represent' other organizations) are brought into equilibrium within a framework of contractual relations" (Jensen and Meckling, 1976, p.289). Under this "nexus of contractual relations" view, the actual boundaries of firms are ambiguous and have less meaning with regard to influencing incentives.

With the introduction of multi-task, principal-agent analysis, Bengt Holmstrom, Paul Milgrom, and John Roberts have reemphasized the importance of firm boundaries in addressing agency problems. When agents are required to complete several tasks, employment within a firm offering muted incentives may be efficient (Holmstrom and Milgrom, 1991). High powered incentives on a particular task that is easy to measure may cause the allocation of effort away from other important but hard to measure tasks. Holmstrom and Roberts (1999) offer an example of a sales operation.

> [S]ales people carry out three tasks: making current sales, cultivating long-term customer satisfaction, and gathering and relaying information on customer needs. If the latter two activities are important and the three activities compete for the agent's time, then the marginal rewards to improved performance on each must be comparable in strength; otherwise, the ill-paid activities will be slighted. Because performance in non-selling activities is arguably hard to measure, it may be best to provide balanced, necessarily low powered incentives for all three activities (Holmstrom and Roberts, 1999, p.164).

When "making current sales" is the only relevant task and is easily measurable, firms can rely on outside sales forces compensated with higher commission rates (Holmstrom and Roberts, 1999). The multi-task approach also has implications for job design and asset ownership. For example, firms providing balanced, low-powered incentives to their in-house sales staff often place limits on outside activities to ensure against the allocation of effort towards outside activities (Holmstrom and Milgrom, 1991). Also, the theory provides "an account of how activities might be grouped, with some employees specializing in tasks that are hard to monitor and others in activities that are easily monitored" (Holmstrom and Milgrom,

1991, p.243). In this case, firms provide strong incentives for employees who perform easily monitored tasks and low-powered incentives for those that perform hard to monitor tasks. Incentives also play a role regarding the efficient use of assets when the value of those assets is partially a function of how they are used. "Under these conditions, when the principal owns the returns from the asset, the optimal incentive contract will provide only muted incentives for the agent to produce output in order to mitigate any abuse of the asset or any substitution of effort away from asset maintenance. However, when the agent owns the asset returns, the optimal incentive contract will provide more intensive incentives to engage in production, in order to alleviate the reverse problem" (Holmstrom and Milgrom, 1991, p.234).

In 1994, Holmstrom and Milgrom built on their multi-task approach in attempting a synthesis of the transaction cost and agency cost theories in *The Firm as an Incentive System*. Here, the authors contend that choices regarding asset ownership, performance incentives, and employer discretion over employee activities are interrelated and co-vary in predictable ways. Within the boundaries of the firm, employers direct and supervise workers who use tools owned by the firm and typically are paid fixed wages. Arm's length contractual relationships "tend to involve purchases from a worker who chooses his or he own methods and hours and owns the tools used and is paid only for quantities supplied" (Holmstrom and Milgrom, 1994, p.972). Firm and market relationships as defined by these three elements represent alternative incentive systems. According to the authors, the choice between these systems is influenced by variations in three exogenous parameters: measurement costs, asset specificity, and uncertainty (Holmstrom and Milgrom, 1994, p.973). Holmstrom and Milgrom return to their industrial sales model in explaining. "[W]hen the cost of measuring sales performance is high (e.g. because it involves team selling) or when hard-to-measure nonselling activities are important, it is more likely that the agent's optimal incentives will conform to the attributes of employment: modest commissions, firm ownership of customers, and no right for the agent to sell products to other manufacturers. On the other hand, when performance is easy to measure or when nonselling activities are unimportant, incentives will conform to attributes of independent representation: strong output-based incentives, customer ownership, and freedom to sell the products of other manufacturers" (Holmstrom and Milgrom , 1994, p.974).

The problem with relying solely on incentive-based theories of economic organization is that they fail to elaborate on important aspects of organizational choices. They do not explain, for example, why competing tasks like making sales, cultivating customer satisfaction, and gathering information should ever be bundled into a single job. From an incentive-only perspective, the sales task (assuming low measurement costs) should be separate and rewarded on an output basis while the other tasks (assuming high measurement costs) can be combined into a different position and compensated by fixed wages. An incentive approach does not describe the productive efficiencies associated with bundling several tasks in one job which often are taken as given. These efficiencies in production are critical determinants of job design and the boundaries of the firm. Their inclusion calls for theoretical approaches that address issues of production and coordination. The following section describes such an approach.

CAPABILITIES THEORY AND THE BOUNDARIES OF THE FIRM

In his seminal contribution, G. B. Richardson (1972) defined capabilities as the "knowledge, experience and skills" appropriate to the performance of productive "activities" that need to be completed and coordinated. Furthermore, he categorized activities according to the types of capabilities they require. Two or more activities are *similar* if they require the same productive capabilities. The resource-based view of the firm beginning, perhaps, with Penrose (1959) has illustrated how excess capacity with respect to a given capability provides incentives to expand production in new—but similar—directions. Such a view provides a productive efficiency rationale for firm diversification by engaging in activities subject to *economies of scope* with knowledge being a shared input in the production of more than one output. Although Richardson defines similarity of activities based on the capabilities they require, activities are *complementary* if they contribute to different stages of a coordinated production process.

Richardson points out that the coordination of productive activities may be provided by intra-firm direction, inter-firm cooperation or through market transactions. Complementary activities may reflect

standardized production such that arm's length, spot market transacting efficiently coordinates the plans of independent producers in different phases of production. Alternatively, the relatively specialized nature of some complementary activities may require close cooperation between firms. Richardson refers to such activities as *closely complementary*. Richardson concludes that complex networks of cooperation exist "because of the need to coordinate closely complementary but dissimilar activities" (Richardson, 1972). Richardson argues that the number of complementary activities undertaken by firms usually is limited because they often are dissimilar.

Richardson's analysis implies that the nature of productive capabilities is an important determinant of firm boundaries. The process of production itself is fraught with uncertainty, not over asset specificity or agency problems, but about how productive knowledge and routines needs to be coordinated (Langlois and Foss, 1999). The coordinating role of capabilities is embedded in the knowledge and routines that constitute production, and may not be analytically separable from the productive activity itself. Such an orientation opens the door to economic models that take firm heterogeneity seriously. Firm heterogeneity derives from the differences in capabilities among firms.

This does not suggest that transaction cost economics (TCE) lacks explanatory power. Rather, as David Teece has noted, "[i]n order to fully develop its capabilities, transaction cost economics must be joined with a theory of knowledge and production (Teece, 1990). TCE's explanatory power is not independent of the productive capabilities at work within the firm. That is, productive capabilities, to some degree, determine the costs of transacting. By holding the degree of asset specificity and agency costs constant and varying the capabilities required in the production process, one sees that the changing nature of capabilities alone provides a rationale for alternative organizational forms.

Given the limits of firms' productive knowledge, skill, and experience, economic change often poses challenges for the existing array of complementary capabilities. In particular, product and process innovations can create "dynamic transaction costs" that require changes in organization for their implementation. Dynamic transaction costs consist of the "costs of persuading, negotiating, coordinating and teaching outside suppliers….[or] the costs of not having the capabilities you need when you need them" (Langlois, 1992, p.113). It is these dynamic transaction costs associated with economic change that may call for changes in vertical relationships within the supply chain. "When the market cannot provide the right capabilities at the right time, vertical integration may result; and when the firm lacks the right capabilities at the right time, vertical disintegration may occur" (Langlois, 1992, p.113) . Systemic innovations which require simultaneous changes in multiple stages of the supply chain may require vertical integration to carry them out. On the other hand, if firms do not possess the requisite capabilities, innovation may lead to vertical disintegration as firms rely on the market to complete the necessary activities. Learning by firms and markets over time creates incentives for altering vertical boundaries as relative capabilities change (Langlois, 1992).

Though TCE and the capabilities view can be construed as complementary, the relationship between the two approaches is an uneasy one, owing to the different ways the two theories have been operationalized. TCE has been exploited as a tool of static optimization. As such, the methodologies it employs reflect the power—and limitations—of neoclassical theory. Economic capabilities, by contrast, derive their explanatory power from plausibility rather than tractability. Empirical support for the economics of capabilities is most often case-study based, and therefore derives from a context of changing markets. Capabilities reflect production as an innovative process not wholly compatible with static optimization.

TEMPORAL SPECIFICITY, DYNAMIC TRANSACTION COSTS, AND THE HOLD-UP PROBLEM

Despite their distinct approaches, there is some degree of overlap between the capabilities and transaction costs literatures. As Langlois and Foss explain, "[c]apabilities would certainly seem to qualify as specific assets – they are specialized to firms; they [often] have low (or no) value in alternative uses…." (Langlois and Foss, 1999, p.234). Broadly speaking, both the capabilities and transaction cost

approaches emphasize the problems of hold-up for market coordination in the face of specialized investments (though the nature and sources of hold-up differs). According to the transaction cost approach, the hold-up problem occurs when "one who makes a relationship-specific investment is vulnerable to a threat by other parties to terminate the relationship. This threat then permits these parties to obtain better terms than were initially agreed" (Milgrom and Roberts, 1992, p.599). It is this threat of *ex post* appropriation of quasi-rents that can discourage *ex ante* decisions to make specific investments in the first place. Under vertical integration, joint profit maximization attenuates incentives for this type of opportunism (Williamson, 1985). This focus on this "hold-up" potential implicitly assumes that problems of production and coordination have been solved *apriori* and therefore vertical integration enhances investment incentives not by affecting the overall value of the economic surplus created by production, but its distribution.

If we take a broader definition of hold-up to include threats that hamper the efficient production of a good or service, a different rationale for vertical integration emerges. In the capabilities literature, hold-up arises from the potential loss of economic surplus associated with not having access to the capabilities you need when you need them. Its focus on coordination problems is essentially Coasian in spirit (Langlois and Foss, 1999). The central issue of organization is not in addressing incentives to haggle over the economic surplus but in creating it in the first place. Here, vertical integration represents a response to the dynamic transaction costs of finding, persuading, negotiating, or teaching outsiders (Langlois, 1995, p.37). If these costs are prohibitive, it may be difficult to access the capabilities on the market necessary to complete a stage of production (or stages of production). For value to be created, firms may need to develop the appropriate capabilities and internalize these stages of production.

Within the transactions costs literature, authors have identified a variety of forms of asset specificity including site specificity, physical asset specificity, human asset specificity, brand name, and dedicated asset specificity. A form of asset specificity that is relevant here is "temporal specificity." As a source of hold up, temporal specificity problems arise when the timing of activities is an important component for efficient production. "Where timely performance is critical, delay becomes a potentially effective strategy for exacting price concessions. Knowing that interruptions at one stage can reverberate throughout the rest of the project, an opportunistic supplier may be tempted to seek a larger share of the gains from trade by threatening to suspend performance at the last minute. Even though the skills and assets necessary to perform the task may be fairly common, the difficulty of identifying and arranging to have an alternative supplier in place on short notice introduces the prospect of strategic holdups" (Masten, Meehan, and Snyder, 1991, p.9). Vertical integration thus represents an organizational response to this hold up potential. The "greater potential for strategic holdups in market transactions should increase the costs of contracting relative to internal organization and thus also increase the probability of integration" (Masten, Meehan, and Snyder, 1991, p.11).

This paper contends that temporal specificity has been miscast as a problem of opportunism rather than one of coordination. The holdup problems associated with temporal specificity often can be more accurately portrayed as ones arising from not having access to the necessary capabilities at the time they are needed. The inability to achieve an efficient coordination of activities through arm's length relationships leads to a reduction in productive efficiency and an attending loss in economic surplus. These circumstances are more likely to occur when the novelty of an innovation causes capabilities not to be widely available on the market. Vertical integration may be necessary to ensure the efficient coordination of activities along the supply chain.

It may be useful to revisit an example from the transaction costs literature in elaborating on this point. An excellent illustration of the empirical analysis applying the concept of temporal specificity is Craig Pirrong's "Contracting Practices in Bulk Shipping Markets: A Transactions Cost Explanation" (Pirrong, 1992). In describing the relationships between carriers and shippers in this market, Pirrong notes the persistence of a wide variety of contractual practices including spot market contracts, forward contracts, medium term contracts, long term contracts, and vertical integration. Pirrong begins his analysis by asking the following question: Why are long-term contracts and vertical integration so prevalent when the traditional sources of transaction costs like physical asset specificity, site specificity, and human asset

specificity are notably absent from shipper-carrier relationships? Ships are mobile (no site specificity), commodity-specific ships typically are numerous (no physical asset specificity), and navigation/ seamanship skills are highly transferable across ships (no human asset specificity). The author answers by arguing that certain conditions cause "time and space factors" that create challenges for spot market and forward contracting relationships between shippers and carriers. Temporal specificity "exists when certain factors create a strong tie between a particular pair of transactors; that is, the value of a transaction between buyer B1 and Seller S1 is more valuable that a transaction between B1 and any other seller or S1 and any other buyer. This tempts traders to engage in wasteful behaviors in order to capture rents" (Pirrong, 1992, p.136).

According to Pirrong, three factors affect the severity of temporal specificity between shippers and carriers. "Any factor that reduces the substitutability of ships as carriers of a particular cargo should increase the quasi rents that make spot contracting costly" (Pirrong, 1992, p.137). Conditions that increase substitutability and thus support spot market contracting include the wide availability of geographically dispersed alternative cargo sources, very thick markets, and the ability to utilize general purpose cargos. The absence of any one of these conditions increases the severity of temporal specificity and thereby undermines spot market relationships. Under these circumstances, shippers and carriers can potentially choose to engage in forward contracting.

> Temporal specificity-created quasi rents are therefore reduced when shippers and carriers negotiate forward contracts rather than spot contracts. By contracting sufficiently far ahead, the shipper protects himself from expropriation of his specialized assets….[B]y contracting [a number of] days ahead the shipper can access [more] ships….Similarly, forward contracting allows carriers to expand their competitive alternatives in order to protect themselves against the unilateral threats of any shipper. These factors make the forward contracting bargaining range smaller than the spot market contracting bargaining range, which reduces the scope for inefficient bargaining and strategic behavior" (Pirrong, 1992, p.137).

While forward contracting increases the number of carriers (shippers) a shipper (carrier) can access and therefore mitigates the severity of temporal specificity, it also comes with its own costs. "Forward contracting requires the parties to enter into agreements prior to the revelation of all relevant information about the value and costs of services for which they have contracted. Incomplete forward contracts restrict the ability of shippers and carriers to respond in a value-maximizing way to unforeseen circumstances. They must weigh the advantages of avoiding exposure to market power by increasing competition against the costs of reduced flexibility The ability of forward contracting to mitigate temporal specificity still does not explain why long term contracts and vertical integration are common in bulk shipping" (Pirrong, 1992, p.4).

Single voyage, forward contracts often suffer from "contractual specificities" because they do not address matching problems for the shipper and carrier after the contract has been completed. "Contractual specificities exist when a contract between a shipper and carrier expires, but all other vessels the shipper can utilize and all the other shippers the carrier can service are under contract for a considerable amount of time thereafter. The shipper and carrier are in a bilateral monopoly situation during the interval between the end of their contract and the termination of the next expiring agreement" (Pirrong, 1992, p.140). The use of forward contracting to negotiate agreements that coordinate starting and expiration dates among multiple carrier-shipper pairs are likely to be prohibitive and greatly suffer from incomplete contract problems. Under these circumstances, long-term contracts or vertical integration may be the only practical response. "If all parties negotiate contracts with the duration equal to the expected economic life of the ship they employ, or if shippers purchase their own vessels, they eliminate their dependence on others, avoid the need for cooperation, and thereby reduce transaction costs" (Pirrong, 1992, p.141).

The traditional holdup story emphasizing ex post threats as a means to appropriate a larger share of quasi rents likely distorts the true nature of the problem. Instead of fearing being vulnerable to

opportunism, shippers (carriers) simply may be uncertain whether they will be able to access the carrier (shipper) they need when they need them. If market conditions are such that temporal specificity problems are severe, there may not be a supplier (or buyer) available at the right place and time to extract those quasi rents to begin with. An example (from a footnote) that Pirrong uses to illustrate the importance of market thickness in supporting spot market exchange is instructive for our purposes here.

> Many travelers arrive at O'Hare Airport in Chicago needing to travel downtown. Knowing this, many cabs congregate there. As a result, cabdrivers know that they will get a fare with high probability, and travelers know that they will get a cab after a short wait (if they have to wait at all). Under these circumstances, a traveler need not contract ahead with a cabbie to assure a prompt pickup; decentralized spot contracting is feasible here. If, however, a rider needs a cab on an isolated side street in an exclusive suburb where few people use cabs, it is highly unlikely that he would readily find one just by standing in front of his house. Since the traffic is thin, few cabs go there without having prearranged a fare. Thus, the suburban traveler will call a cab company – that is contract forward – to obtain a ride (Pirrong, 1992, p.138).

Note that the suburban traveler's concern is not in being gouged by an opportunistic cabdriver who can extract a greater share of the rents when alternatives are not readily available but in having access to cab service at all. Thus the story with regard to temporal specificity is really one of having access to capabilities at the appropriate time. In the context of systemic innovations, the market for required capabilities can be thin indeed for reasons already discussed. Here, forward contracts simply won't do because those capabilities do not exist. Vertical integration then represents a response to the lack of capabilities on the market which fundamentally is a problem of coordination, not opportunism. The next section provides a case study illustrating the importance of capabilities in explaining the boundaries of the firm.

A TALE OF TWO SUPPLY CHAINS: LI & FUNG AND ZARA INTERNATIONAL, INC.

Background

A number of different retail models account for the bulk of new clothing sales in the United States. Department stores provide a large selection of brand names and in-house lines. Recently, several department stores have added celebrity sponsored merchandise and exclusive collections from well-known designers. "Kohls has clothes from Vera Wang and MTV reality TEV star Lauren Conrad. Macy's lineup includes Madonna, while J.C. Penney Co. has merchandise from Mary-Kate and Ashley Olson" (Holmes and Zimmerman, 2011, p.B2). Specialty stores like the Gap Inc. and Aeropostale, Inc. offer limited product lines "but great depth within those lines" (Pride and Farrel, 2008, p.468). Discount stores like Walmart and Target sell both brand name and private brands at low prices with Target also offering collections from well-known designers at a discount. Fast fashion retailers like Zara, H & M, and Forever 21 have emerged as a new retail model during the last ten years. Zara, for example, makes "the latest styles available to customers shortly after they appear on the catwalk or are worn by a trend setting celebrity. Zara excels in getting 'must have' merchandise to the market quickly at a cost low enough that consumers feel they can look for something new every week or two" (Klammer, 2009, p.2). Despite the diversity of approaches, relatively few players are responsible for the lion's share of sales activity in all retail categories. "In the United States, the top five chains came to account for more than half of apparel sales in the course of the 1990s, and concentration levels elsewhere, while lower, also rose during the decade. Increased concentration was generally accompanied by the displacement of independent sores by retail chains, a trend that had also helped increase average store size over time" (Ghemawat and Nueno, 2003, p.3).

The variety of retail strategies and the heterogeneity of product lines employed by apparel companies have led to diverse sourcing strategies and the adoption of a wide array of organizational forms. Supply

chains in the apparel industry require the coordination of the following activities: apparel design, materials sourcing, manufacturing, distribution, and retail. Many of the traits that characterize apparel production (high labor intensity, low skill labor, low asset specificity, and limited economics of scale) imply that decentralized methods of market coordination should predominate (Dulbecco and Vagneron, 2001). In organizing their supply chains, apparel companies face a number of challenges that make timely performance a critical factor in determining success. "The world market for clothes and textiles experiences wide, cyclical and unpredictable variations in the volume of demand, the quality of the products, and the production mix" (Dulbecco and Vagneron, 2001, p.58). Traditional supply chains often have had difficulty in responding efficiently to rapid changes in the marketplace. "Historically, fashion collections have been seasonal and new designs would hit the market only four times a year… Most firms lock in more than 50% of their production prior to the start of a season, hoping they guessed right on fashion trends." (Klammer, 2009, p.1). Off the mark predictions of consumers' buying patterns often result in large amounts of inventory that have to be sold at sharply discounted prices at end-of-season clearance sales.

The emergence of new information technologies in the 1980s and 1990s spurred retailers, brand name manufacturers, and specialized intermediaries to reconfigure their supply chains to reduce cycle times and allow for quicker responses to changes in consumer demand. The adoption of "Computer-aided design" (CAD) and "Computer-aided manufacturing (CAM) has led to significant efficiencies in coordinating the design and manufacturing stages of production (Collier and Collier, 1990). Radio Frequency Identification (RFID) technology has also played an important role in facilitating these changes. RFID tags included small integrated circuits that were used for product identification. "Powered by the radio waves of the reader, the RFID tag sends back data to the reader wirelessly. This replaces clumsier bar codes for product identification, making it possible to more easily monitor the movement of a boxful of products and even individual products throughout the supply chain" (Fung, Fung, and Wind, 2008, 59). Information technologies such as CAD, CAM, RFID and the development of industry-specific computer software packages promoted "quick response (QR), a set of policies and practices targeted at improving coordination between retailing and manufacturing so as to increase the speed and flexibility of responses to market shifts… QR had led to significant compression of cycle times enabled by improvements in information technology and encouraged by shorter fashion cycles and deeper markdowns, particularly in women's wear" (Ghemawat and Nueno, 2003, p.3).

The advances in information technologies have had differential implications for the organization of supply chains; in some cases, they promoted the vertical specialization of production while in others, they prompted vertical integration. We offer case studies of two very different organizational responses: the management of decentralized, vertically specialized supply chains by Li & Fung and the vertically integrated supply chain of fast fashion innovator, Zara, Inc. These alternative approaches reflect the different strategic imperatives of their respective retail models. As an upstream intermediary, Li & Fung is responsible for handling apparel orders for retailers such as The Limited, Target, Kohls, Wal Mart and others. The company specializes in managing the coordination of all phases of production among a broad and internationally dispersed network of independent suppliers. As a pioneer in fast fashion, Zara offers high-fashion concept apparel and accessories for relatively low prices with new designs being shipped to stores recurrently throughout the year. Zara relies heavily on its own in-house design, production, and distribution capabilities in executing its strategy.

Using the capabilities framework, this paper compares these two very different production strategies. It takes the approach recommended by Langlois and Foss (1999) by assuming "that all incentive problems can be eliminated by assumption and concentrate on coordination and production-cost issues only" (Langlois and Foss, p.206). We attempt to show that the adopted organizational forms represent efficient responses to potential holdup problems caused by coordination delays in production rather than threat of opportunistic behavior by trading partners. Li & Fung's development of a thick network of independent suppliers at all stages of production ensured its access to the capabilities it needed when it needed them. Zara accomplishes this by following a vertical integration strategy for its trendier, highly time-sensitive

apparel items. The different approaches reflect the strategic demands of their two very different retail models.

Li & Fung and Network Orchestration

Created during the first decade of the twentieth century in Southern China, Li & Fung "was the first Chinese-owned export company at a time when the China trade was controlled by foreign commercial houses" (Magretta, 1998, p.192). The ability of its members to speak both Chinese and English was an important competitive advantage for this small family firm. In response to declining margins during the 1970s, Li & Fung transitioned from essentially a brokerage service to a regional sourcing agent with its geographical reach extending to Taiwan, Korea, and Singapore. The company, in that capacity, would find sellers of specific items requested by its clients. With its wide and deep knowledge of suppliers over East Asia, Li & Fung moved into managing and delivering manufacturing programs. In response to a buyer's request for a certain end product, Li & Fung sources the inputs, contracts production with individual factories, and monitors quality and delivery. Over time, Li & Fung has evolved into a leader in global supply-chain management (Magretta, 1998).

Li & Fung describes what it currently does as network orchestration. It draws upon the capabilities of many geographically dispersed suppliers within their network. "The network orchestrator designs the overall supply chain, drawing together multiple factories in different regions to collaborate on a single product" (Fung, Fung, and Wind, 2008, p.3). Li & Fung is responsible for coordinating production among these many dispersed manufacturers along with ensuring quality and timely delivery. The example below illustrates the network orchestration process.

> On May 30, a U.S. retailer places an order with Li & Fung for 300,000 pairs of men's twill cargo shorts. Li & Fung owns no factories, no weaving machines, no dye, no cloth, no zippers. It does not directly employ a single seamstress. Yet one month later, the order is shipped. [T]he buttons come from China; the zippers come from Japan; the yarn is spun in Pakistan, and woven into fabric and dyed in China; and the garment is sewn together in Bangladesh. Because the customer needs quick delivery, the order is divided among three factories. Yet every pair of shorts has to look as if it were made in one factory. If the order had come in two weeks later, it would have resulted in a completely different supply chain, using different partners from a network of 8,300 suppliers around the globe. Like a message routed through the Internet, the project moves among the best specific path chosen from a broader network (Fung, Fung, and Wind, 2008, p.10).

Li and Fung's organizational structure consists of 170 divisions which are responsible for implementing the network orchestration process. Most divisions handle one major customer while others serve "a group of customers with a similar need" (Magretta, 1998, p.5). Having each division focus on a single customer facilitates the integration of operations and aligns incentives between Li and Fung and its buyers. Though the suppliers are independent, Li & Fung maintains control "through educating, rigorous monitoring, and testing to ensure compliance. The supply chain is loosely linked, but compliance systems are tight" (Fung, Fung, and Wind, 2008, pp.72-73). To create incentives for suppliers to invest in these relationship-specific compliance systems and maintain quality standards, Li & Fung tries to ensure that 30 to 70% of an individual supplier's business comes from Li & Fung (Magretta, 1998).

In describing Li and Fung's ability to coordinate the diverse capabilities involved in the production of apparel, it is useful to consider the different types of knowledge that are required in the production of a product. Rebecca Henderson and Kim Clark (1990) have developed a conceptual framework that describes the two types of knowledge needed in product development. "First, it requires component knowledge, or knowledge about each of the core design concepts and the way they are implemented in a particular component. Second, it requires architectural knowledge or knowledge about the ways in which the components are integrated and linked together into a coherent whole" (Henderson and Clark, 1990, p.11). Li & Fung has reduced the costs of qualitative coordination between suppliers by choosing modular

product designs that allow for changes at the component level without altering a product's underlying architecture. It does this "by pulling apart the manufacturing value chain and optimizing each step" (Magretta, 1998, p.104). "The modularization of the manufacturing process meant that different parts of the manufacturing process could be handled in pieces and coordinated across factories" (Fung, Fung, and Wind, 2008, p.7).

This type of product design promotes autonomous innovation which allows for changes of an individual component of a product without requiring changes in other components or the product's architecture (Langlois and Robertson, 1995). For example, if Li and Fung decides to switch to different buttons in completing an order for the men's twill cargo pants, it can do so without having to alter the fabric, dye, zippers or any other component that goes into the production of cargo pants. Cargo pants as a category conveys the product's architecture which determines how individual components are put together. The modular product design of apparel like cargo pants within a large network allows Li & Fung to tap the diverse competences of its many suppliers. "This broader network allows the firm to connect to the capabilities it needs (or those its customers need) wherever they are in the world. Whereas competitive advantage once was defined largely by the capabilities the company directly developed and owned, now it depends on the capabilities inside as well as those capabilities outside that the company can connect to" (Fung, Fung, and Wind, 2008, p.17).

Network orchestration combined with the production of modular products allowed Li and Fung to better synchronize its supply decisions with changing market demand. It is able to do this by implementing a strategy of delayed differentiation or postponement. "Using delayed differentiation, a firm delays or postpones the final customization of a related bundle of products (and/or the shipment of a product to geographical markets) to the extent possible, pending more accurate product- and market-specific demand information" (Anand and Girota, 2007, p.697). This final product can be modified in ways that more accurately reflect consumer demand by postponing "the addition of the components that differentiate the product until the later stages of the process" (Feitzinger and Lee, 1997, p.117). Li and Fung refer to this process as "flexible commitments." "Flexible commitments can lock up the capacity of a supplier without specifying the final design and color until the last moment. The trust between Li and Fung and its supplier network means that it can reserve undyed yarn from its yarn supplier. This locks capacity at the mills for the weaving and the dyeing, with the promise that they get an order of a specific size but no details. Five weeks before delivery, Li and Fung lets the supplier know which colors to use and when additional fabric and trim will be delivered. The factory then has three weeks to deliver the final garments. This orchestration gives suppliers predictability, while providing the buyers with rapid and accurate response" (Fung, Fung, and Wind, 2008, p.58). The ability to delay the revelation of all attributes of an order grants Li & Fung the ability to respond to changes in market demand on a timely basis, thereby reducing the amount of clearance sale inventory for clients at the retail level. The modular nature of production in which autonomous changes can be made at the component level gives network orchestrators the flexibility to make last minute decisions without disrupting the overall production process.

With modular products, network orchestrators can engage in forward contracting with their suppliers without incurring the costs of a reduction in flexibility. Unlike the bulk shipping market story, the incomplete nature of these forward contracts actually increases flexibility by allowing network orchestrators to place orders in advance without revealing all the attributes of that order until just before production begins. Through forward contracting, network orchestrators have access to a larger number of suppliers than what would have been available to them under spot contracting. Under spot contracting, many suppliers may be engaged in contracts with other buyers thereby reducing the number of available alternatives. Network orchestrators thus can retrieve a greater array of capabilities by contracting in advance with suppliers. It is the ability to access capabilities within the network on a timely basis (rather than avoiding opportunism) that represents the primary motivation for forward contracting. Additionally, incomplete forward contracts within a modular production system increases flexibility leading to reductions in product cycle times. These factors combined with a very thick network of suppliers largely accounts for Li and Fung's success.

Zara International, Inc. and Fast Fashion

Zara, Inc. is a recognized leader in fast fashion (aka cheap chic), producing and selling trendy clothing at affordable prices. It currently operates 4,600 stores world-wide that specialize in the sale of women, men, and children's apparel for the under forty crowd. In order to appeal to fashion conscious consumers, Zara chooses store locations that are "upscale in prime high street areas such a as the Champ Elysees in Paris, Regent Street in London and Lexington Avenue in New York City" (Fraiman and Singh, 2002, p.7). Its stores receive twice a week deliveries of fresh designs from Zara's central warehouse in La Coruna, Spain; Zara typically makes 20,000 products available each year, a number four times greater than its competitors (Surowiecki, 2004, 193; Caro and Gallien, 2010, p.257). Zara has demonstrated an ability to respond quickly to changes in market demand; it is "able to originate a design and have finished goods in stores within four to five weeks in the case of entirely new designs and two weeks for modifications (or restocking) of existing products" (Ghemawat and Nueno, 2003, p.9). For example, Zara was "able to tilt the in-store inventory from equestrian themes to black within two weeks of the September 11 terrorist attacks" (Fraiman and Singh, 2002, p.6).

Zara's focus on fashionable but inexpensive apparel and its ability to respond quickly to changes in consumers' buying patterns at the regional level has been at the core of its success. Its "enhanced design capabilities result in products that are of greater value to consumers and hence elicit a greater willingness to pay by charging higher prices on 'trendy' products than on more conservative products" (Cachon and Swinney, 2011, p.779). Its capabilities also have increased profitability in other ways. "Shipping fewer pieces, in a great variety of styles more often requires shorter lead times and high-level flexibility. As a consequence of offering fewer pieces more often, fast fashion retailers collect larger percentages of the full price and thus achieve higher net margins on sales" (Tokatli, 2008, p.23). Only 15-20% of Zara's sales are at discounted prices; for competitors, these numbers range from 30-40% of sales (Caron and Gallien, 2010). The small batch sizes combined with the frequent introduction of fresh products creates a contrived scarcity that induces shoppers to make repeat visits so they will not miss out on the latest designs (Cohen, 2007). The typical Zara customer visits Zara stores 17 times per year compared to 3 - 4 times for the competition (Framian and Singh, 2002, p.7). In addition, this strategy "helps Zara reduce its exposure to fashion faux pas. The company produces batches of clothing in such small quantities that even if it brings out a design that no one will buy -- which happened during an unseasonably warm autumn in 2003 – it can cut its losses quickly and move on to another trend" (Tiplady, 2006). In addition, Zara spends very little on advertising, relying heavily on word-of-mouth. Zara's strong financial performance attests to the success of its efforts, earning an income-to-sales ratio of 14% in 2010 (Inditex Group Annual Report, 2010).

Interestingly, Zara's creation was the result of a failed business involving Inditex, then a Spanish-based manufacturer of women's lingerie products. "In 1975, after a German customer cancelled a sizeable order the firm opened its firsts Zara retail shop in La Coruna. The original intent was simply to have an outlet for cancelled orders but the experience taught the firm the importance of a marriage between manufacturing and retailing – a lesson that has guided the evolution of the company ever since" (Ferdows, Lewis, and Machuca, 2003, p.62). It is this recognition of the need for close coordination that led Inditex to integrate forward into retailing. In addition to its flagship brand Zara, Inditex added the following smaller retail chains: Massimo Dutti, Pull & Bear, Bershka, Stradivarius, and Oysho (Ghemawat and Nueno, 2003). During the 1980s, Inditex increased Zara's reach by opening stores in all the major cities in Spain. In the 1990s, it focused on aggressively expanding internationally. In cooperation with Toyota, Zara "began to make major investments in manufacturing logistics and IT, including the establishment of a just-in-time manufacturing system, a 130,000 square meter warehouse close to corporate headquarters in Artexio, outside La Coruna, and advanced telecommunications system to connect headquarters and supply, production, and sales locations" (Ghemawat and Nueno, 2003, p.9). While most retailers rely on relatively standardized software packages in organizing their supply chains, Zara internally developed its own unique logistical system.

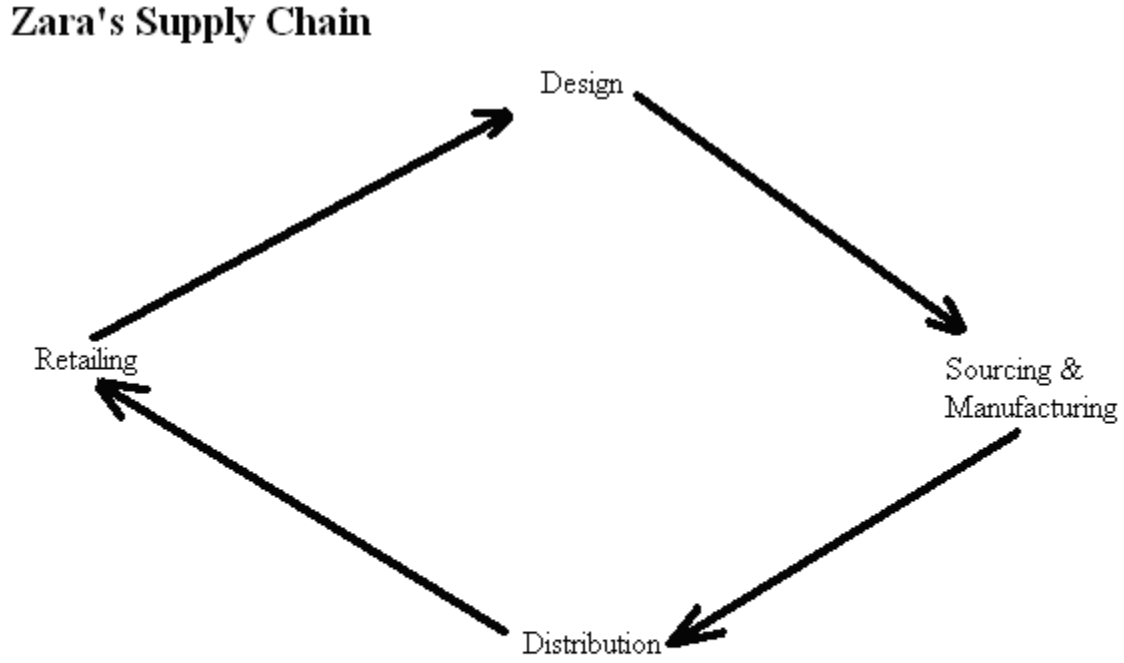
Gerard Cachon and Robert Swinney have characterized fast fashion systems as combining two core components: "short production and distribution lead times, enabling a close matching of supply with

uncertain demand" (quick response techniques) and "highly fashionable product design" (enhanced design techniques) (Cachon and Swinney, 2011, p.778). Its focus on enhanced design sets Zara apart from other supply chains (like Li and Fung) that make use of postponement (or delayed differentiation) in implementing their quick response techniques. "The distinction between postponement and enhanced design is one of degree. Postponement creates variants from a base model (e.g., different color panels for the same phone), whereas enhanced design creates significantly different product variants from component inventory (e.g., a skirt or dress slacks from the same material)" (Cachon and Sweeney, 2011, 780). Enhanced designs tend to be more complicated and unique because they are shaped by the "most attractive and promising trends spotted at fashion shows and by cues taken from mainstream consumers" (Tokatli, 2008, p.21). Applying the conceptual framework of Henderson and Clark (1990), enhanced design "changes a product's architecture but leaves the components, and the core design concepts that they embody, unchanged" (Henderson and Clark, 1990, p.12). Enhanced design implies the creation of new product architecture for each new design. Zara's competitive advantage lies in its ability to quickly create these new products in response to shifting consumer preferences across the diverse markets it serves.

In organizing its global operations, Zara has adopted a hybrid model that "uses a rooted [vertically integrated] network for the more complicated and time-sensitive products – like women's suits in seasonal colors – and a footloose [outsourced] model for the simpler and predictable items, like men's shirts in classic colors" (Ferdows, 2009, pp.145-146). In-house production tends to be reserved for trendy items that require enhanced designs and have short shelf lives while more traditional items are outsourced to independent suppliers. "Zara's enviable accomplishment is in keeping these networks focused on different strategic targets: the footloose network on reducing production costs and filling temporary and seasonal capacity gaps, and the rooted network on developing unique production capabilities that support its fast-response supply chain system" (Ferdows, 2009, p.146).

It is useful to think of Zara's supply chain as a loop rather than a linear linking of complementary functions with a beginning and end (see the exhibit below). Zara's success in pursuing its fast fashion strategy depends upon the efficient retrieval and communication of customer information in synchronizing activities along the supply chain. The strong feedback mechanisms embedded in Zara's organizational structure increase Zara's responsiveness to changes in buying patterns. "Zara's supply chain is organized to transfer both hard data and anecdotal information quickly and easily from shoppers to designers and production staff. It's also set up to track materials and products in real time every step of the way, including inventory on display in the stores. The goal is to close the information loop between the end users and upstream operations of design, procurement, production, and distribution as quickly and directly as possible" (Ferdows, Lewis, and Machuca, 2004, p.2).

FIGURE 1



Source: Ghemawat and Nueno

Whether outsourced or manufactured in-house, Zara's commercial teams of designers, store specialists, and buyers (300 in total) are responsible for designing all of Zara's products at its corporate headquarters in La Coruna (Ferdrow, Lewis, and Machua, 2003, p.64). Each commercial team specializes in one product line: men's, women's, or children's apparel. Their personnel keep close track of consumer preferences at the retail level on a real time basis. Within each commercial team at headquarters, store specialists are responsible for communicating with retail store managers within a particular geographic region. "Each store specialist was likely to have retailing experience and was chosen for her commercial design sense and feel for market trends because it was the job of store specialists to feed market information from the stores back into the design and production decisionmaking" (Fraiman and Singh, 2002, p.5). In addition to information at the retail level, design choices were influenced by high fashion, popular media, trend spotters, and the internet (Ghemawat and Neuno, 2003, p.10). "While designers were catwalk-influenced and expected to adapt haute couture style for the mass market, 'they are not themselves encouraged to be ivory tower aesthetes making distinctive fashion statements'" (Fraiman and Singh, 2002, p.2). Designers allocated their product development efforts between products for the current season based upon the latest information and on the product mix for the initial collection of the following season (Ghemawat and Nueno, 2003, p.10). Once designs have been created using drawings and the CAD system, in-house workers manually translate those designs into clothing samples. After further collaboration, members of these commercial teams then decide which designs will go into production.

With regard to production, the "make or buy decisions are usually made by the procurement and production planners. The key criteria for making this decision are required levels of speed and expertise,

cost-effectiveness and availability of sufficient capacity" (Ferdows, Lewis, and Machua, 2003, p.64). In general, Zara produced about 50% of its products in-house relying on twenty of its own factories to handle the highly fashionable, time-sensitive apparel items. "More basic items that were more price-sensitive than time-sensitive were particularly likely to be outsourced to Asia, since production in Europe was typically 15% - 20% more expensive for Zara" (Ghemawat and Nueno, 2003, p.11). Zara sources its fabric and most inputs from external suppliers, often relying on Comiditel, an Inditex subsidiary that specialized in purchasing materials for all Inditex divisions. "Over half of these fabrics are purchased undyed to allow faster response to mid-season color changes" (Ferdows, Lewis, and Machua, 2003, pp.64-65). Comiditel also "managed the dyeing, patterning, and finishing of gray fabric for all of Inditex's chains, not just Zara" (Ghemawat and Nueno, p.11)… After procuring fabric, in-house production called for the cutting of fabric according to the specifications of the design. "Zara's factories were heavily automated, specialized by garment type, and focused on the capital-intensive parts of the production process – pattern design and cutting – as well as on final finishing and inspection" (Ghemawat and Nueno, 2003, p.11). Zara then relies on hundreds of local subcontractors to perform sewing operations, a relatively low-skilled, labor-intensive activity with limited economies of scale. "Deliveries between Zara factories and subcontractors occurred many times a week with subcontractors picking up new work as they left off completed work. Overall turnaround time for sewing ran a week or two. Assuming it had the fabric in stock, Zara was in a position, with its in-house design, pattern-making and cutting capabilities and its network of sewing subcontractors, to go from start to finish on a style production within as little as ten days" (Fraiman and Singh, 2002, p.6).

After the completion of the sewing, finishing and inspection operations, the garments were packaged in labeled plastic bags and then sent to one of Zara's central distribution centers. "Every item has to go through one of Zara's three logistical hubs, all in Spain. A pair of pants made in China for Zara is shipped to Spain first, even if the final destinations is China" (Bjork, 2011, p.B2). The distribution center in Artexio, for example, "operated on a dual shift basis and featured a mobile tracking system that docked hanging garments in the appropriate bar coded area and carrousels capable of handling 45,000 folded garments per hour" (Ghemawat and Nueno, 2003, p.11). The primary role of these hubs is the coordination of production with retail with most items staying in the distribution center for less than a day. The distribution center received orders from each retail store twice a week via hand held computer with the filling of the order subject to the center's approval based upon availability and past sales performance for items in short supply (Ghemawat and Nueno, 2003, p.12) . "Orders for each store are packed into separate boxes and racks (for hanging items) and are typically ready for shipments 8 hours after they have been received" (Ferdows, Lewis, and Machua, 2003, p.65). Shipments to individual stores occurred on a regularly scheduled basis. They "were made out of the distribution center twice a week by truck to Europe and by airfreight to stores outside of Europe so that stores received goods within 24 – 36 hours of shipments in Europe and 1 – 2 days outside Europe" (Fraiman and Singh, 2002, p.6).

In order to attract wealthier, fashion conscious consumers, Zara placed its stores in "highly visible locations, often including the premier shopping streets in a local market and upscale shopping centers" (Ghemawat and Nueno, 2003, p.14). Its stores tended to be relatively uniform with respect to size, window displays, lighting and overall layout. "The uncluttered arrangement of goods in uncrowded spaces coordinated by color made the experience of shopping more like that in high-end luxury stores, and quite different from that offered by 'value' marketers" (Fraiman and Singh, 2001, p.7). The retail store managers play an important role in ensuring that stores have the right products at the right time. "In addition to overseeing in-store personnel, store managers decided which merchandise to order and which to discontinue and also transmitted customer data and their own sense of inflection points to Zara's design teams. In particular, they provided the creative teams with a sense of latent demand for new products that could not be captured through an automated sales tracking system" (Ghemawat and Nueno, 2003, p.14). Slow moving merchandise is quickly removed from store shelves and is often shipped to other stores or back to the central warehouse (Ferdows, Lewis, and Machua, 2003, p.66). Zara's vertically integrated production structure does not imply a reliance on centralized decision-making; its processes allowed Zara to tap the local knowledge of store managers in making critical design choices.

Zara's decision to vertically integrate many of the activities involved in the production of highly fashionable, time sensitive items is based upon its need to have the appropriate capabilities at the right time in meeting its strategic objectives. For Zara, efficient execution takes priority over cost minimization. "At Zara, rapid timing and synchronicity are paramount. To this end, the company indulges in an approach that can best be characterized as 'penny foolish, pound wise.' It spends money on anything that helps to increase and enforce the speed and responsiveness of the chain as a whole" (Ferdrows, Lewis, and Machua, 2004, p.2). Zara purposely maintains excess capacity in its operations so it can have the capabilities available to respond quickly to changes in the market. "For Zara, in-season replenishment did not require incremental capacity to be found or much higher costs incurred at the last minute; instead the close-to-sale time manufacturing permitted an ongoing reallocation of resources in the ordinary course with minimal disruption. In-house manufacturing capacity had been reserved and was available, but exactly what was to be manufactured could be determined within a few weeks of when the garments would appear for sale in the stores" (Fraiman and Singh, 2002, p.6). Zara's factories normally worked on a single shift basis while "the distribution center operates four days per week with the precise number of shifts depending on the volume of products that have to be distributed" (Ferdows, Lewis, and Machua, 2004, p.65). The high fixed costs of maintaining reserve capacity and the limited scale of batch production translates into high unit costs. Zara, however, was able to recoup these "costs on the garments through markup because people will pay a premium for the right garment at the right time" (Fraiman and Singh, 2002, p.7).

Zara has not been monolithic in its pursuit of a vertical integration strategy in producing its trendier products. When capabilities were available on the market and could be accessed on a timely and cost effective basis, Zara contracted with independent suppliers to provide those capabilities. Zara (or more accurately Inditex) purchased all its fabric and many other inputs from external suppliers. The close proximity of a large number of small workshops allows Zara to rely on outsiders to perform sewing operations. Access to the right capabilities at the right time ultimately was the motivating factor in the make-or-buy decision.

DISCUSSION AND CONCLUSIONS

The different supply chains adopted by Li & Fung and Zara reflected the capabilities that each needed in pursuing their individual strategies. A strategy based on enhanced design, for example, poses serious challenges for the type of quick response techniques implemented by Li & Fung. Li & Fung's ability to delay differentiation depends upon the modular production of apparel with well-established product architectures. "The requisite know-how to produce a commodity product is usually highly codified and easy to transfer from one factory to another, inside or outside of the company" (Ferdows, 2009, p.144). Differentiation does not arise from the creation of new architectural knowledge as in the case of enhanced design but results from the modification of an existing architecture with the addition of components late in the production process. This allows for Li & Fung to simultaneously engage in forward contracting with producers and differentiate their products. These modular product designs, however, tend to "accelerate the process of turning the product into a commodity, resulting in smaller profit margins…" (Ferdows, 2009, p.142). Thus profitability depends as much on minimizing production costs as it does on responding to consumers' demands. Li & Fung's success largely is the result of its ability to access a broad array of capabilities from a thick network of low cost suppliers.

By contrast, Zara's vertical integration strategy is extraordinarily costly, entailing the maintenance of reserve capacity to flexibly respond quickly to changes in market demand. Its focus on enhanced design requires the capability to create new product architectures that reflect the latest consumer trends and that can be translated quickly into saleable merchandise. The necessity for the continual creation of new architectural knowledge presents significant problems for market coordination. Forward contracting with suppliers is impractical because of the impossibility of relying on modular product designs and the ability to predict in advance consumer tastes. Spot contracting with suppliers on a real time basis is risky because many of them may be contractually engaged with other buyers. In addition, the nature of architectural

knowledge is not codified but is largely tacit and therefore difficult to communicate quickly to outsiders (Henderson and Clark, 1990). Zara's internal design decisions depend upon an organizational structure that fosters the close cooperation of its designers, market specialists, and retail managers. Because of the high fixed costs of maintaining these capabilities, Zara's profitability is a function of producing trendy apparel tailored to local tastes at the right time and in the right quantities. Zara's customers have accommodated such a strategy, demonstrating a willingness to pay higher prices for their unique merchandise.

The focus here on capabilities does not suggest that incentives are not important; in certain cases incentives may be the primary determinants of firm boundaries (Masten and Snyder, 1991, pp.20-1). It is also important to note that vertical integration does not necessarily imply low-powered incentives for employees. For example, Zara's store "managers received a fixed salary plus variable compensation based primarily on their store's performance with the variable component representing up to one-half of the total, which made their compensation very incentive-intensive. Since prices were fixed centrally, the store managers' energies were primarily focused on volume and mix" (Ghemawat and Nueno, 2003, p.15). While time sensitivity and fashion styling were the most critical factors, relative production cost was an important consideration in the decision to outsource production. "Inditex owned 21 Zara factories, each of which were separately managed. Factory managers' bids were assessed against third party supplier bids to make sure that in-house manufacturing stayed competitive" (Fraiman and Singh, 2002, p.5).

The capabilities approach also stresses that the boundaries of firms are not static. As the level of capabilities changes both within firms and across markets, firm boundaries will adjust in response to changes in relative efficiencies between internal and market coordination. For example, there is evidence to suggest that Zara has quietly begun to outsource the production of some of its trendy and time-sensitive products to firms in distant countries. "As a number of supplier firms in countries such as Morocco, India, and Turkey have gained the competence to manufacture intricately worked high quality garments with the required flexibility and speed, Zara has turned to sourcing from these countries. It appears that instead of Zara changing the geography of jobs, the geography of competencies has changed Zara" (Tokatli, 2008, p.21).

REFERENCES

Anand, Krishnan and Girota, Karan. (2007). "The Strategic Perils of Delayed Differentiation," *Management Science*, May, 53(5): 697-712.

Aghion, Philippe and Holden, Richard. (2011). "Incomplete Contracts and the Theory of the Firm: What Have We Learned over the Past 25 Years?" *Journal of Economic Perspectives*, Spring, 25(2): 181-197.

Bjork, Christopher. (2011). "Cheap Chic Apparel Sellers Heat Up U.S. Rivalry on Web," *Wall Street Journal*, September 6, pp. B1 & B2.

Caro, Felipe and Gallien, Jeremie. (2010). "Inventory Management of a Fast-fashion Retail Network," *Operations Research*, March-April, 58(2): 257-273.

Cohen, M.L. (2007). "Zara International, Inc." International Directory of Company Histories. Retrieved December 10, 2010 from Encyclopedia.com: http://www.encyclopedial.com/doc/lG2-3479900105.html.

Collier, Billie and Collier, John. (1990). "CAD/CAM in the Textile and Apparel Industry," *Clothing and Textiles Research Journal*, March, 8(3): 7-13.

Dulbecco, Philippe and Vagneron, Isabelle. (2001). "Competition, Co-operation and Subcontracting: Lessons from the Clothing Industry in Thailand," *The European Journal of Development Research*, 13(2): 58-80.

Feitzinger, Edward and Hau, Lee. (1997). "Mass Customization at Hewlett-Packard: The Power of Postponement," *Harvard Business Review*, 75(1): 116-121.

Ferdows, Kasra. (2009). "Shaping Global Operations," *Globalization, Competitiveness, and Governability*, 2009, 3(1): 136-148.

Ferdows, Kasra., Lewis, Michael., and Machua, Jose. (2003). "Case Study: Zara," *Supply Chain Forum*, 4(2): 62-67.

_____. (2004). "Rapid-fire Fulfillment," *Harvard Business Review*, November, Reprint R0411G, pp. 1-7.

Fung, Victor., Fung, William., and Wind, Yoram. (2008). Competing in a Flat World: Building Enterprises for a Borderless World. U.S.A.: Wharton School Publishing.

Fraiman, Nelson and Singh, Medini. (2002). "Zara," *Columbia Business School Case Study*.

Ghemawat, Pankaj and Nueno, Jose. (2003). "Zara: Fast Fashion." *Harvard Business School*, Case 703-497, April 1: 1 - 35.

Heath, Chip and Staudenmayer, Nancy. (2000). "Coordination Neglect: How Lay Theories of Organizing Complicate Coordination in Organizations," *Research in Organizational Behavior*, Vol. 22, pp. 155-193.

Henderson, Rebecca and Clark, Kim. (1990). "Architectural Innovation: The Reconfiguration of Existing Product Technologies and the Failure of Established Firms," *Administrative Science Quarterly*, March, 35(1): 9-30.

Holmes, Elizabeth and Zimmerman, Ann. (2011). "Shedding Dowdy for Cool, Department Stores Revive," *Wall Street Journal*, August 4th, B1-2.

Holmstrom, Bengt and Milgrom, Paul. (1994). "The Firm as an Incentive System," *American Economic Review*, Sept., 84(4): 972-991.

Holmstrom, Bengt and Roberts, John. (1998). "The Boundaries of the Firm Revisited," *Journal of Economic Perspectives*, 12(4): 73-94.

Inditex Group. (2010). "Brief of Annual Report," Retrieved on September 14, 2011 from http://www.inditex.es/en/shareholders_and_invstors/investor_relations/annual_reports/brief

Jensen, M., and Meckling, W. (1976). "Theory of the Firm: Managerial Behavior, Agency Costs, and Ownership Structure," *Journal of Financial Economics* 3: 305-360.

Klammer, Tom. (2009). "Competing with Fast Fashion at Zara," *IMA Educational Case Journal*, September, Art. 1, 2(3): pp. 1-3.

Klein, B., Crawford, R., and Alchian, A. (1978). "Vertical Integration, Appropriable Rents, and the Competitive Contracting Process," *Journal of Law and Economics* 21(2): pp. 297-326.

Langlois, Richard N. (1992). "Transaction-cost Economics in Real Time," *Industrial and Corporate Change*, 1, 1: 99-127.

_____. (2009). "Economic Institutions and the Boundaries of the Firm: The Case of Business Groups," *University of Connecticut*, *Department of Economics Working Paper Series*, Working Paper 2009-23.

Langlois, R. and Foss, N. (1999). "Capabilities and Governance: The Rebirth of Production in the Theory of Economic Organization" *Kyklos* 52(2): 201-218.

Magretta, Joan. (1998). "Fast, global, and entrepreneurial: supply chain management, Hong Kong style," *Harvard Business Review*, Sept.-Oct., 76(5): 102-125.

Masten, Scott., Meehan, James., and Snyder, Edward. (1991). "The Costs of Organization," *Journal of Law, Economics and Organization*, Vol. 7, pp. 1-25.

Monteverde, Kirk and Teece, David. (1982). "Supplier Switching Costs and Vertical Integration in the Automobile Industry," *Bell Journal of Economics* 13: 206-213.

Penrose, Edith. (1959). *The Theory of the Growth of the Firm*. Oxford: Basil Blackwell.

Pride, William and Ferrell, O.C. (2008). *Marketing*. Boston: Houghton Mifflin Company.

Richardson, G. B. (1972). "The Organization of Industry," *The Economic Journal*, v.82, n.327, pp. 893-896.

Tipaldy, Rachel. (2006). "Zara: Taking the Lead in Fast-Fashion," *Bloomberg Businessweek*, Apr. 4, p. 1.

Tokatli, Nebahat. (2008). "Global Sourcing: Insights from the Global Clothing Industry – the Case of Zara, a Fast Fashion Retailer," *Journal of Economic Geography,* 8, pp. 21-38.

Williamson, O. (1985). *The Economic Institutions of Capitalism*. New York: The Free Press.