

Postponement: an inter-organisational perspective

Biao Yang, Ying Yang, J Wijngaard

▶ To cite this version:

Biao Yang, Ying Yang, J Wijngaard. Postponement: an inter-organisational perspective. International Journal of Production Research, 2007, 45 (04), pp.971-988. 10.1080/00207540600698886. hal-00512909

HAL Id: hal-00512909 https://hal.science/hal-00512909

Submitted on 1 Sep 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Postponement: an inter-organisational perspective

Journal:	International Journal of Production Research
Manuscript ID:	TPRS-2006-IJPR-0065
Manuscript Type:	Original Manuscript
Date Submitted by the Author:	26-Jan-2006
Complete List of Authors:	Yang, Biao; University of Ulster, School of Business Organisation and Management Yang, Ying; University of Wales, School of Management and Business Wijngaard, J; University of Groningen, Department of Production Management
Keywords:	POSTPONEMENT, CAPACITY PLANNING, SUPPLY CHAIN MANAGEMENT
Keywords (user):	organisational structure



Re-submitted to International Journal of Production Research

Postponement: an inter-organisational perspective

Biao Yang^{*}, Ying Yang and Jacob Wijngaard

* Corresponding author Dr. Biao Yang School of Business Organisation and Management Ulster University at Jordanstown Co. Antrim BT37 0QB Northern Ireland Email: b.yang@ulster.ac.uk Tel. + 44 (0) 2890366359

Dr. Ying Yang School of Management and Business Cledwyn Building, Penglais Campus University of Wales Aberystwyth, SY23 3DD United Kingdom

P P Professor Jacob Wijngaard Department of Production Management Faculty of Management and Organization University of Groningen PO Box 800, 9700 AV Groningen The Netherlands

Postponement: an inter-organisational perspective

Keyword(s): postponement, organisational structure, capacity planning

Abstract: In view of the slow rate of postponement applications, this paper attempts to examine postponement strategies from an inter-organisational perspective. The paper first reviews the literature on different postponement strategies (including logistics postponement, production postponement, purchasing postponement and product development postponement) mainly within the context of supply chain or supply network. Then, two research questions are formulated. This is followed by the investigation into the intra-organisational structure and capacity planning in postponement applications. The paper concludes with some suggestions for further research.

1. Introduction

In the current business environment, there is an increasing tendency towards the outsourcing and globalisation of different types of operations. Many manufacturers are shifting production and sourcing to countries with lower labour costs often far from primary consumer markets. Some of them have even started to evaluate whether they must continue to assemble products themselves or whether they can outsource production entirely (Liker and Choi, 2004). The extension of production and distribution across diverse national boundaries does not merely reflect the development of global supply chains but also new orientation of companies towards core competencies. This reflects the idea of networks of relationships (Harland, 1996), where business organizations are moving to a borderless inter-dependence business

Page 3 of 35

International Journal of Production Research

environment. Consequently, a company can not develop or exploit its own resources except in conjunction with of others. The optimisation of such a networked system is determined by the effectiveness of supply chain relationships as well as delivery performance and total supply chain costs. This paper contributes to the understanding of this issue by focusing on an increasingly supply chain- and network- oriented concept – postponement.

Postponement may be defined as a dimension of sequence and timing based on the concept of substitutability (Bucklin, 1965), subsequently maintaining the opportunity of interchangeability and irreversibility. The benefits of postponement mainly come from changing the sequence of activities or delaying them in time: (1). Changing the sequence of activities could lead to a demand aggregation effect where the planning is based on aggregated demand (Zinn, 1990). The demand aggregation effect reduces the forecasting horizon, resulting in better resource planning and allocation. It also reduces risk by pooling the variance of the demand, which is aligned with the concept of centralisation of inventories (Maister, 1976; Eppen, 1979). (2). Delaying activities in time enables companies to learn from the behaviour of the demand and other environmental factors (Aviv and Federgruen, 2001). This provides an opportunity to incorporate more actual information (e.g. about customer demands) into companies' decision-making processes (Lee, 1998). The forecast can also be improved after observing some actual demand (Fisher and Raman, 1996). In this regard, postponement may be viewed as an approach for synchronising the supply chain (Goldratt, 1990). It introduces time buffers at those points where the lack of information might destroy the synchronised flow. That is, it provides more time for those processes which require further information. Much has been written in the

literature on how to implement postponement by re-designing product architectures and/or manufacturing processes. However, the implementation of postponement strategies also affects the structure of the supply chain, since postponed activities will be most likely to be placed closer to the time and locus of consumption. Consequently, these effects can be reflected in the design of supply chains from product development, to sourcing and on to final distribution of products. In recent times, the development of a postponement strategy has been managed from an integral supply chain perspective (Feitzinger and Lee, 1997; Pagh and Cooper, 1998; Yang and Burns, 2003). This view opens up more opportunities for postponement applications, since a company may make up a delay in the speed of one of the supply chain areas (resulting from a postponement strategy) by improving the others.

Now that the need for a higher level of postponement across supply chains is increasingly accepted, the challenge of implementation becomes critical. Despite the conceptual underpinnings of postponement, its applications are still not as widespread as expected. Surveying more than 350 supply chain professionals at mid-sized and large companies, Oracle Corporation (2004) finds that postponement is an underutilised concept. In a more recent questionnaire survey across the food, clothing, electronics and automotive industries, Yang *et al.* (2005) find most managers expect postponement to be less used in three years. They further reveal that most significant difficulties in postponement implementation are related to how a company manages its external networks (suppliers or customers). This concurs with Waller *et al.* (2000) who argue that, as an inter-organisational concept, postponement has received little attention. To address this, this paper attempts to investigate postponement strategies in the context of supply chains or supply networks. The paper precedes by reviewing the

Page 5 of 35

literature on postponement mainly from an inter-organisational perspective. Then, two research questions are formulated. This is followed by the investigation into the interorganisational structure and capacity planning in postponement applications. The paper concludes with some suggestions for further research.

2. Literature Review

Postponement, once only a strategy for differentiating a product, has matured into a way of global thinking in product design, production, logistics and marketing (Bowersox and Morash, 1989; Van Hoek, 2001). A main distinctive principle of postponement is to obtain more actual information in order to define and translate the customer's needs into a concrete product or service specification. Four types of postponement strategies can be formulated based on the state of inventories (see figure 1): (1). Logistics postponement (finished goods): It seeks postponement opportunities in the final movement of products, which have taken their final form in advance of customer orders, to the customer. This is facilitated by the emergence of ecommerce, where virtual inventories are independent on the inventories' physical location at the time orders are placed. Internet retailers can thus manage inventory to fulfil customer orders by postponing the location of inventory to upstream supply chain echelons (their suppliers) until the arrival of customer orders (Bailey and Rabinovich, 2005). In practice, Amazon.com usually chooses to work closely with its vendors and the United States Postal Service to ensure that it can use such a postponement strategy (drop-shipping) to handle the volume and delivery timing of a popular product; (2). Production postponement (work-in-process): Keeping undifferentiated semi-finished products for as long as possible will increase a company's flexibility in response to changes in customer demand. In addition, the

company can reduce supply chain cost by keeping undifferentiated inventories, e.g. on a basis of material profiles (Umble and Srikanth, 1990; Yang and Burns, 2003); (3) Purchasing postponement (components or raw materials): This strategy allows companies to postpone the purchase of incoming components or raw materials until demand is known, eliminating the risk of holding obsolete inventory in stock. The strategy, together with production postponement, allows companies to avoid building up inventories of finished goods in anticipation of future orders. In parallel with this postponement strategy, the practice of vendor managed inventories (VMI) calls for the vendors to take responsibility for replenishing the retailer's customers, thus delaying the ownership of inventory; and (4). Product development postponement (no inventory): One such case is the Drive Test system from Agilent Technologies, which measures the coverage, effectiveness and quality of wireless phone base stations (Elliott and Hughes, 2000). They face the challenge of bringing together three different sets of rapidly changing knowledge to provide a valuable test solution for cell manufacturers and operators. For example, new and existing wireless communication standards emerge rapidly. Traditionally, they have focused on maximising resource utilisation over all parts of the project while often not properly co-ordinating different knowledge streams. The controller has to be re-developed later on due to the very latest information. Following the postponement philosophy, they assign most resources to the riskiest and longest lead-time knowledge development first, and postpone low-risk subprojects as late as possible so that all pieces of the total knowledge package come together simultaneously. The lower risk subprojects also take into account the latest knowledge from high-risk subprojects, suppliers, customers and other sources to minimise integration risks. The dotted line in Figure 1 reflects how postponement is associated with the customer decoupling point (CODP),

 where in the supply chain the customer order penetrates and that distinguishes forecast and order-driven activities.

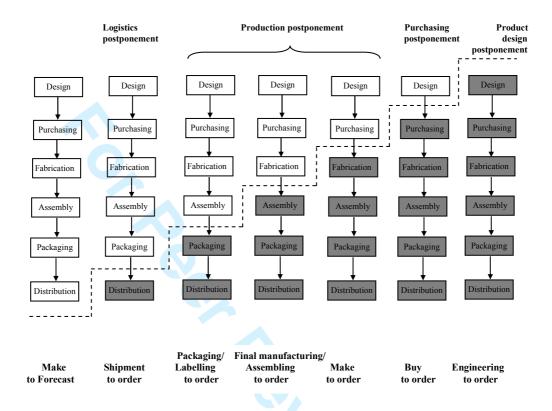


Figure 1. Postponement strategies and positioning the customer order decoupling point (adapted from Yang and Burns 2003)

Too often, postponed activities may not happen within an organisation within unified ownership. Success of a postponement strategy increasingly relies on a supply chain wide thinking. For example, production and purchasing postponement should be exploited to reduce inventory levels throughout the supply chain rather than simply shift inventory burden to the lower tiers of the chain. This is attributed both to shorter forecast cycles and shifting inventory upstream the supply chain to a less expensive generic state. Therefore, in postponement applications, downstream players must be concerned about possible order fulfilment problems occurring upstream in their supply chains. Otherwise, downstream players order as late as possible to avoid the lack of flexibility associated with giving an order at an early stage, while their

suppliers do not have the opportunity to better plan and schedule their production and distribution (e.g. even out operation load), and the utilisation of their resources system drops to the disadvantage of the performance of the whole supply chain. In a supply chain setting, postponement draws attention also because it is relevant to the positioning the customer order decoupling point (Yang and Burns, 2003), as mentioned earlier. Similarly, postponement may contribute to a glocalisation strategy, where the organizational type can be centralised upstream whereas a decentralized approach is for downstream activities focusing on delivering a customized product on time. In the apparel-manufacturing industry, Jin (2004) also exploits how apply postponement to achieve benefits of both global and domestic sourcing methods, thus reducing inventory risk arising from market uncertainties. However, this stream of research mainly focuses on the spatial/functional dimension of the supply chain and fails to address its organisational dynamics. Determining the right level of postponement depends not only on appropriate analysis of customer requirements but also existing organisational capabilities. Here we pose the following research question: what organisational structures may facilitate the implementation of postponement in an inter-organisational context?

Further, as demonstrated in our classification of postponement previously, postponement may be considered as a way of substituting additional capacity for investment in inventory. Postponement delays the irreversible commitment of capacity to certain activities until the last possible moment. It may, to some extent, improve the utilisation of capacity through the effective re-allocation of assets or resources to the appropriation locations in the supply chain (Van Hoek *et al.*, 1999). However, optimisation models and techniques for resource allocation and capacity utilisation

Page 9 of 35

under conditions of uncertainty, bottleneck and congestion are typically encountered in postponement strategies. Postponement highlights a need for greater customercentral activities, thus drastically reducing the possibility to optimise the process in general (Olhager, 2003). Given the variability of customer demands, this implies a reduction in the utilisation of resources. For example, production postponement may require short delivery times, changeovers increase and production batch decrease to provide more product variety over shorter periods. It may also require more resources to provide quick throughput times, since throughput can only be satisfied by production resources and not inventory of final goods. Too often the implementation of postponement requires certain levels of capacity and resources to be reserved for those activities that benefit most from the additional information gained by the delay. Capability of the postponed transformation process to respond to high demand variability requires excess capacity and high throughput efficiency (Skipworth and Harrison, 2004). This requires creating and maintaining slack resources, thus placing enormous demands on capacity planning. That is, a challenge arises regarding how to determine optimal resource requirements to support a postponement strategy.

3 Organisation Structure

It has been suggested in the literature that postponement is relevant to networked and modular organisational forms (Van Hoek *et al.*, 1998 and Yang *et al.*, 2004a). Van Hoek *et al.* (1998) reveal that the implementation of postponement strategies may be key to the creation of a network organization, which can in turn compete on all three dimensions of global competition: global efficiency, local responsiveness and worldwide learning. In fact, the principle of postponement generally provides a basis for understanding the evolution of different channel structures. In comparing the internationalization literature and the network literature, Ford (2002) notes that an international business has its roots in speculative production and the need to export or dispose of existing production overseas to achieve economies of scale or to generate a return on development expenditure, while a network organization is concerned with many new types of companies entering the network to provide a wider range of offerings, relating much more closely to postponement and flexibility in production. Note that the accelerated infusion of e-commerce (e.g. e-marketplaces) into business operations is drastically changing the way businesses operate. In e-commerce, buyers and suppliers are offered a chance to easily find a partner for spot purchases. The resultant more visibility implies more potential partners in a less stable network. To complicate matters further, switching costs in e-commerce are likely to be lower than they are for traditional ways of doing business. The buyer can often switch suppliers with just a few mouse clicks. Therefore, postponement generally reflects the push towards the idea of companies as networked organisations in the current business environment, where profits are increasingly gained from collaborations both within an organisation and across organisational borders.

Considering the role of product and production modularisation in a successful postponement strategy, modular organisations also come no surprise in postponement (for a thorough discussion see Yang *et al.*, 2004a). A recent development on postponement research - exploiting postponement in response to uncertainty (Yang *et al.*, 2004b), also brings its relevance and interest to modular organisations. Uncertainty created by increasing heterogeneity in customer demands and technology complexity is often cited as a primary factor driving the use of modular organisational forms (Baldwin and Clark, 1997 and Schilling and Steensma, 2001). In literature,

 modular organisations are characterised by different elements like dynamic teaming (Tu et al., 2004), dynamic community (Galunic and Eisenhardt, 2001), and contract manufacturing, alternative work arrangement and alliance (Schilling and Steensma, 2001). Obviously, the above organisational structures all have potential of recombining their resources in a variety of configurations to rapidly respond to changing business conditions (Galunic and Eisenhardt, 2001). Therefore, they should not merely positively link to postponement, but all other strategies which rely on gaining access to necessary components through strategic alliances and outsourcing. Here, we consider more specific inter-organisational dimensions in different postponement strategies. To facilitate our discussion, we take a manufacturer perspective in the subsequent subsections. 3.1 Product development and purchasing postponement

With a purchasing postponement strategy, little inventory exists to cushion production or scheduling problems. Raw materials or components must be replenished quickly and arrive where and when they are needed. Manufacturers thus need to realize the potential benefit and importance of strategic and cooperative buyer-supplier relationship. That is, purchasing postponement emerges as manufacturers experimented with strategic partnerships with their immediate suppliers. Some manufacturers go even farther to taking advantage of the immediate supplier's capability and technology during the product design stage through early supplier involvement. This is relevant to product development postponement, which provides an opportunity to reduce design lead times and costly redesigns. By initially focusing on product specification decisions that are likely to remain stable, design decisions

about less stable portions of the product can be postponed until better information is available (Yang *et al.*, 2004b).

 In the literature, much has been written regarding how improving supply management provides organisational coordination required in the implementation of postponement. In the automotive industry, for example, Van Hoek and Weken (1998) investigate the effects of postponed manufacturing in the supply chain where suppliers manage part of the inbound complexity associated with material flows (by providing standardized modules) while the final modular assembly is extended into distribution centres and even customer sites. This development is likely to require suppliers to determine the changes and strategies necessary to accommodate production operations within a modular context. Consequently, suppliers need to acquire new technological knowhow and more strategic supply chain management activities. This also underscores the need for a complete re-shaping of supply networks. It is the interactions and interrelationships among supply chain players that enable the successful exploitation of the postponement strategy. Postponement requires the adoption of more collaborative working practices based on the integration of business processes within a wide community of business partners. At this point, postponement cannot be successfully leveraged without considering the organizational relationships.

Among those, formalisation is a key dimension to capture an inter-organisational structure. In an inter-organisational context, formalisation refers to the degree to which the supply chain and network is controlled by explicit rules, procedures and norms that collectively prescribe the rights and obligations of the individual companies (Choi and Hong, 2002). It may require the development of well-defined

 interfaces (e.g. through rules, procedures, norms and values) between organisation units. Formalisation is closely associated with explicit and codified standardisation (Mintzberg, 1979). Standardization of business processes is necessary to allow communication and integration between business partners within the supply network due to the complexity of processes in the supply chain (Gunasekaran et al., 2001). Here we propose that the implicit yet clearly understood norms lead to flexibility required for product development postponement, while the explicit and formalized procedures and rules regarding the tight restriction on the magnitude and frequency of allowable quantity changes are positively to the implementation of purchasing postponement. In its implementation of product development postponement, Toyota relied on its suppliers in their product development: its suppliers (designers) deeply and simultaneously explored broad regions of the design space, and provided a set of possibilities, while Toyota continuously compiled information on the market and technological developments, various constraint and trade-offs between the alternatives (Ward et al., 1995). This reflects the long-term based manufacturer-supplier relationship involving Toyota's suppliers' investment in developing ideas and plans for a new product before the design is finalised. That is, product development postponement calls for a clear understanding necessary for the continuation of the relationship. This is also in line with the exiting literature that addresses explicitness as the characteristic of formalisation associated with clarity and fairness. Furthermore, product development postponement leads the decision (on critical dimensions about the vehicle) to delay until the last possible moment, ensuring that its suppliers have deeply and simultaneously explored broad regions of the design space. To this end, Toyota continued compiling information on the environment, the market, technological developments expanding the known feasible region, and various

constraint and trade-offs between alternatives, gradually narrowing the set of possibilities and converging slowly toward a single solution (Ward *et al.*, 1995). This implies that implicitness is also necessary for flexibility necessary to make changes to design specifications.

A manufacturer can also choose its level of purchasing postponement, which is the choice of the decoupling point (work-in-process or finished goods) at which the manufacturer responds to a specific customer request. Meanwhile, the manufacturer needs to provide advanced warning to a supplier for a change to an order quantity's size or timing, but the warning may come within the effective lead time horizon, leaving the supplier to somehow cope with the change. This is reflective of not only the capabilities of the manufacturing process but also the nature of the supply contract (Krajewski et al., 2005). Flexible supply contracts and thus relatively frequent schedule revisions require a supplier to frequently deal with changes in production volumes and delivery schedules. It has been well documented in the postponement literature that the above practice (i.e. massive reductions in inventories and the pushing of responsibility and cost further up the supply chain) may have significant consequences for supply chain performance. In the automotive industry, Sako et al. (1994) observe that purchasing postponement (JIT deliveries) have resulted in a shift in inventories from the assembly plant to the supplier in a number of cases. This also underscores the need for a complete re-shaping of supply networks. By contrast, the supply contract restrictions (as a manifestation of formalisation in the supply chain and network) on allowable changes to the timing and size of component orders allow those suppliers to stabilise their manufacturing environments by improving the accuracy of the production and shipping information. Such an environment allows

these suppliers to better time the manufacture of their components with the buyers need for them, thereby reducing the need for finished goods inventories and reducing inventory costs (Krajewski *et al.*, 2005). The tighter restrictions on changes to order quantities improves production schedules, and consequently leads to less safety stocks often moving the decoupling point between process inventories and buyer demands closer to the raw materials level. It can thus be proposed that restrictive supply contracts and infrequent schedule revisions are positively related to the implementation of purchasing postponement.

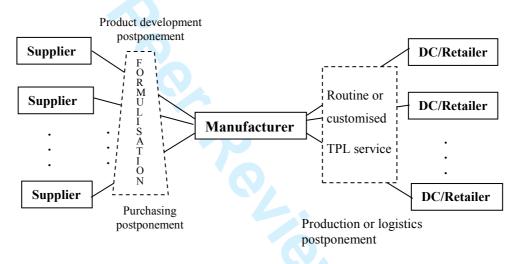


Figure 2. An inter-organisational structure for postponement strategies

3.2 Production and logistics postponement

Postponement can also be extended (from production) further downstream in the delaying of order delivery. For example, if an order's delivery could be postponed to a later time and the truck is not loaded to capacity, the order can be held for later delivery in the hope of consolidating it with other orders that might arrive in the future. At this point, production and logistics postponement may be involved in production, logistics and marketing. However, the need for accurate management of manufacturing-logistics and marketing interfaces is likely to be a bottleneck in achieving the potential benefits of postponement strategies (Van Hoek *et al.*, 1998).

Along the same line, Daugherty and Pittman (1995) state that potential advantages from speedy manufacturing are diminished if the merchandise gets tangled in the distribution chain, since the customer's primary concern is how long it takes to receive the merchandise following order placement.

Here we focus on the relationship between postponement and service strategies for third-party logistics (TPL) providers. Based on complexity of service and type of customer relationship, service relationships can be distinguished between routine TPL service and customised service. This classification is also characterised by economy of scale and scope, coordination and co-operation of operations with the TPL provider (Bask, 2001). For example, customised TPL services demand closer relationships between manufactures and TPL providers than routine TPL services. The manufacturing postponement matches with customized logistics services where final/light manufacturing, together with after-sales services, can be carried out by a third-party logistics provider, e.g. final assembly and packaging of a computer, printer and manual by country. Standard logistics services may include some easy customized operations e.g. centralised warehousing. This matches with logistics postponement which takes the advantages of a centralised inventory, from where finalised products are distributed to a larger geographical area. In this case, inventories can be reduced, and the advantage is high in-stock availability. Similar to this, another emerging practice is virtual warehousing with the development of ecommerce (Matthews and Hendrickson, 2002), where companies may actually hold no physical inventory of supplies or parts. They contract with partners or third-party vendors, who manage an overall level of inventory. The contract companies can centrally manage similar inventories of products for many companies at once. In

Page 17 of 35

 summary, tailoring logistics service to production and logistics postponement need to be responsive, aiming at supply chain flexibility and speed (Bask, 2001).

4. Capacity planning

A major aspect of capacity management is the control of actual and potential bottlenecks (Blackstone, 1989). Based on the assumption of bottlenecks in obtaining information, the drum-buffer-rope (DBR) in theory of constraint (Goldratt, 1990) can be used to explain why postponement should be applied. The DBR is about dealing with bottleneck: The drum is the bottleneck constraint and its pace should be the pace of the entire system; The buffer is inventory held in front of the bottleneck constraint to avoid idleness on the bottleneck constraint; and the rope is a prioritisation mechanism which pulls materials towards the bottleneck. Any time lost at the bottleneck is the time that can never be made up. Traditionally, the supply chain run a rate based on forecast to push the mix of (semi-) finished product downstream the supply chain. However, with constantly changeable demands this cannot improve the effective system throughput. Conversely, it only increases the finished product inventory (running high risk of obsolescence) or reduce necessary throughput (running high risk of out-of-stock). With the growth in demand uncertainty, supply chains now need to strategically locate inventory and capacity to enable both timely and efficient delivery in line with market demands. That is, the most important bottleneck is the market demand. This constraint determines the effective throughput for the supply chain. At this point, postponement enables activities to be performed based on the actual market demands, thus synchronising the entire system to the real demands. It can therefore improve the coordination of resources and the optimisation of activities across the supply chain (Pagh and Cooper, 1998; Yang and Burns, 2003).

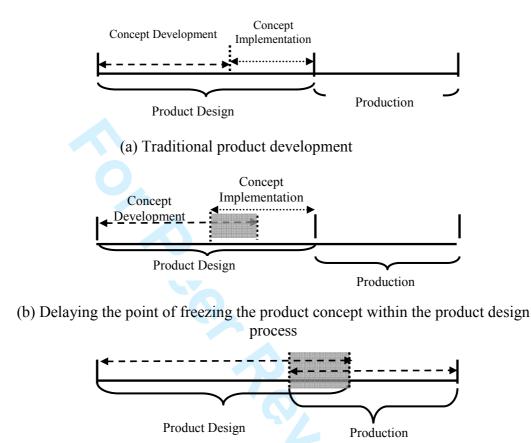
In merging the concept of de-coupling and the customisation of the product, the concept of postponement has further been discussed. Translated into capacity planning, postponement is a risk-based sequencing strategy whereby low-risk activities employ speculative capacity and higher-risk activities are postponed until additional market information has been available (Fisher *et al.*, 1994). The speculate capacity should be made the furthest in advance in order to have greater capacity available for postponed activities.

4.1 The Customer Order Decoupling Point

The decision options on capacity planning can be divided into those trying to modify demand to match the production constraints, and those modifying supply to match the market demand. Olhager et al. (2001) note that, when discussing capacity planning strategies, it is of great importance to distinguish different market environments reflected by different positioning of the customer order decoupling point (CODP). As illustrated in figure 1, the degree of postponement relates to the rational behind the choice of CODP (Hoekstra and Romme, 1992; Olhager, 2003), where product variety usually increases significantly in the chain. The necessary organisational changes (result from different positioning of CODP) have been extensively discussed in the literature (e.g. see Lampel and Mintzberg, 1996; Gilmore and Pine, 1997; Duray et al., 2000). This is also relevant to the concept of P: D ratio (Shingo, 1981), where p represents the production lead-time and D represents the delivery lead-time. In this sense, postponement facilitates capacity planning by utilising the opportunity to stabilise the process upstream the supply chain and invest in a responsive capability downstream in response to market demand variation. This can be highlighted by the role of postponement in shifting the position of CODP.

In Lee's (1998) pull postponement, it is proposed to make the decoupling point earlier in the process so that the differentiation tasks can be delayed to the point when more information about the actual demands is available. Here, the driving force for shifting CODP upstream is to increase the knowledge of the contents of customer orders at the time of production (Olhager, 2003). In a study in the Japanese bicycle industry, Kotha (1996) reveals that capacity planning in pull postponement can be appropriately conducted through the interaction between mass production and mass customization. By the development of the supply network it is also possible to move CODP upstream. This can help to reduce the size of the order specific inventory significantly, but will increase the demands for precision throughout the specification processes including selling, order management, product configuration and service. By contrast, in their definition of leagility, Naylor et al. (1999) utilise postponement to move the CODP downstream and increase the efficiency and the effectiveness of the supply chain. Investment in (sometimes additional) capacity to protect material flow rather than inventory is central to the postponement strategy. To optimise the bottleneck operations, it may also need to have the bottleneck upstream the CODP, rather than using them to respond to volatile demands. The bottleneck capacity can thus be run based on forecast, instead of being postponed (Olhager, 2003; Yang and Burns, 2003). If capacity is constrained it is beneficial to hold stock before the bottleneck, which can be built during slack periods. In practice, companies may separate different strategies in time or place. They may separate out early and late production runs based on the predicted level of uncertainty (Gattorna and Walters, 1996). For products with highly seasonal demand, for instance, a company may manufacture some products to stock in

periods with low demand in anticipation of peak demand (Olhager, 2003). Production is thus levelled and plant utilisation increases.



(c) Delaying some product design activities after production starts

Figure 3. Product design postponement

However, when linking postponement strategies to different supply chain strategies (also see figure 1), it is implicitly assumed that product design, production and distribution takes place in sequence, even in engineering to order. This might not always be true in practice. In their discussion of product development postponement, Yang *et al.* (2004b) also note that, in extremely uncertain markets, companies may not be able to build the perfect product and then get it just right. They may have to get the product to market quickly and then attain feedback from real users. As such, companies can modify the new product idea early in their product development better to meet customers' needs. This implies production activities are initiated while some

 engineering is still going on. Figure 3 further illustrates this overlap. When applied within the product design process, product development postponement is in line with its definition proposed by Yang *et al.* (2004b), as illustrated in figure 3 (b). When further extending it to include the production process in figure 3 (c), the product development postponement strategy may not match any supply chain strategies which treat product design activities as something finished before any production take places. Such a postponement strategy is somewhat analogous to current engineering addressing coordination and integration of functionally different product and production activities. This also reflects Wikner and Rudberg's (2005) view that the increasing pressure on lead-times and customisation highlights the need for some product design activities to be carried out in combination with production activities. It should be noted that a combined product design and production perspective pose a challenge to involve resources from production in the product development process (Wikner and Rudbert, 2005).

4.2 Reactive Capacity Planning

From a manufacturing strategy perspective, postponement is preferable in that the overall demands (the aggregate items number) are relatively easy to accurately forecast. This can be translated in to a capacity requirement plan in terms of aggregate resources. In this sense, postponement reduces costs by allowing some capacity to be held at average demand rather than peak demand. It often allows companies to separate the customization of products from the primary or basic manufacturing of generic (semi-finished) products. This separation also frees primary manufacturing to focus more on large economical runs, while secondary or final manufacturing can be focused on responding to customer wishes. Following Fisher *et al.* (1994) we divide

production capacity into two distinct parts: speculative product capacity (employed before the observation of early indicators of market demand) and reactive production capacity (employed after early demand indicators are observed). We focus on the latter for a postponement strategy.

A manufacturer can simply take the option to stay with a fixed capacity that is capable of handling peak demand or unexpected increases in requirements (Fisher and Ittner, 1999). For example, reactive capacity is set at peak demand, while speculative can be set at average demand. The rational is that the greater its reactive capacity, the greater the proportion of production it can reserve until early demand indicators have reduced demand uncertainty. Companies like Zara intentionally have extra capacity on hand to increase the supply chain's responsiveness to new and fluctuating demands while using postponement to gain more speed and flexibility (Ferdows et al., 2004). By tolerating lower capacity utilisation in its factories and distribution centres, Zara can increase the supply chain's responsiveness to peak or unexpected demands. Zara can thus carry less inventory than its rivals like Benetton, H&M, and Gap (Ferdows et al., 2004). Because the company can sell its products just a few days after they're made, it can operate with negative working capita. The cash thus freed up helps offset the investment in extra capacity. Off-peak periods with excess capacity provide an opportunity to rest, plan for the peak, train new employees and accomplish required maintenance. Additionally, building slack capacity can improve a company's ability to recover from supply chain disruptions quickly, which is of growing interest and importance both from an academic and a practitioner perspective.

 However, excess and thus underutilized capacity hurts financial performance particularly at soft demand. To respond to actual customer orders within short delivery lead times, the literature suggests numerous ways of planning reactive capacity like using flexible working hours (including working overtimes and enforcing working days off) and employing contract and temporary workers. A manufacturer can also lower excess capacity risks by making existing capacity more flexible, e.g. training workers for different tasks. Flexibility is a form of pooling that allows use of the same capacity for a variety of products, which decreases the total capacity required. It may enable some of speculative capacity to become reactive, as orders begin filter in. At Sport Obermeyer, it also reduces reactive production capacity by delaying the reactive production commitment (Fisher et al., 1994). More specifically, it focuses on the timing of capacity changes, e.g. by securing market information earlier. To support a production postponement strategy, the company keeps raw materials and factory-production capacity undifferentiated as long as possible. For example, in addition to warehousing raw materials, the company books factory capacity for the peak production periods well in advance but does not specify the exact styles to be manufactured until a later date.

Further, the manufacturer can exploit external sources (e.g. by contract manufacturing) to augment reactive capacity, without changing the internal production rates. Beyond outsourcing, Hagel III (2002) proposes a leveraged growth strategy tapping into a broad existing base of assets of other companies for the benefit of a company. Under this strategy, Li & Fung can stay responsive to future technological and market shifts while owning none of the facilities involved in processing the raw material into finished goods. As an orchestrator, Li & Fung manages the process network, e.g.

 dividing the process into finer and finer segments, having additional providers populate each segment, and facilitating the collaboration of different companies. To minimise disruption, Li & Fung can re-configure the network (e.g. by calling in another service provider) to get the process back on track. Within a week of the 9/11 event, for example, it shifted production out of facilities based in potentially unstable countries to more secure plants to avoid disruption in supply (Hagel III, 2002). The constant reconfiguring of processes also ensures that the overall network operates with optimum productivity and responsiveness.

A manufacture should also seek to identify opportunities for smoothing the peaks and valleys through active management of demand, e.g. taking active steps to deflect some of peak demand to other time slots at the same facilities. It can develop pricing strategies to encourage its business customers to place their orders earlier, where customers might only need to specify the total volume requirements early. Relevant to this is a commonly practiced method in service - yield management. It creates a partial priority queue by segmenting customers by their willingness to pay and, then, allocating capacity by segment (Kimes, 1989). Yield management also allows overselling of capacity (Biyalogorsky et al., 1999) when demand is uncertain. It gives the seller the potential to make very profitable buyback transactions during the spot period. When capacity is insufficient, sellers can buyback capacity at lower prices than the expected spot price. In the same vein, to solve the problem of temporary capacity shortage, some manufactures may arrange a capacity exchange, through which production capacity (originally assigned to prescheduled orders) is transferred to the rush orders in question (Huang et al., 2005). The promised premium associated with the capacity exchange may be particularly attractive for the urgent capacity

demanders. Meanwhile, original capacity holders may be willing to release the capacity (through delaying their prescheduled orders) in exchange for compensation, which comes from either premiums paid by the capacity demander or allowances provided by the manufacturer (Huang *et al.*, 2005). Van Mieghem and Dada (1999) also argue that before going for production postponement, it may be worthwhile to consider flexible ex-post pricing, which typically falls under marketing's responsibilities. More specifically, they recommend the deferral of the pricing decision until demand uncertainty resolved. For example, a traditional automobile dealer must decide how many cars to buy and hold on its lot before market demand is known and then must negotiate a price with each customer (Van Mieghem and Dada, 1999). They also show ex-post price flexibility can make optimal capacity decisions more insensitive to uncertainty than production postponement.

5. Conclusion and Further Research

The result of postponement implementation in the current business environment is often outsourcing operations and building partnerships and alliances that foster the networking of relationships. A successful application of postponement thus requires agreements that strengthen buyer–supplier cooperation in supply chain and network. To meet the challenges of a demanding marketplace and the need for more costeffective use of resources it is also imperative that a better way be found to balance supply and demand in postponement. In this context, this paper has highlighted two issues in a successful implementation of postponement to address the slow rate of postponement applications: inter-organisational organisational structure and capacity planning. As companies begin to take on the appearance of networks embedded in larger networks, postponement helps companies rethink and overhaul the

 configuration of relationship networks (beyond product and production design). We have also considered different strategies in production capacity planning for a postponement strategy. Any company that chooses to implement postponement should obviously tailor them to its own situation. This study can also be of value to companies focusing on the CODP strategies. Empirical studies would complement and add to the conceptual work in this study. In addition, this paper has only examined the formalisation dimension of supply chain and network from a manufacturing perspective. Further research could be focused on other elements such as centralisation and complexity. To avoid dysfunctional organisational behaviour from the perspective of channel power influence on postponement decisions taken by a dominant player, there is also a need to develop appropriate indicators to monitor the postponement performance in supply chain and network.

In Section 2, the scope of postponement strategies in the existing literature has been discussed including production design, purchasing, production and logistics. The increasing focus on service quality (particularly in the rapidly growing e-commerce) may open up a research opportunity to apply postponement into a service context beyond the above scope. The attributes commonly used to distinguish from physical goods and production activities, such as intangibility, inseparability and perishability, lead to special challenges to design the structure of the service process (Sasser, 1976; Hill *et al.*, 2002). For example, inseparability (i.e. there is a high degree of producer-customer interaction in the production of service) and perishability (i.e. services cannot be inventoried) create uncertainty for managers about the process's time, the product's quality and the facility's accommodation of the consumer's needs. These differences, however, may increase the importance of applying postponement in

 service, similar to just-in-time in service (Duclos *et al.*, 1995). This is also echoed by Van Hoek's (2001) view on the customisation element of postponement. To this end, substantial work is needed to develop and apply the postponement concept to the service process.

In an e-commerce environment, direct contact with suppliers and customers and provision of direct enhancement of information-based services are often used to customise service offerings (Peppers and Rogers, 1997; Bitner, et al., 2000). However, since the Internet is an open system, many companies appear to offer almost identical services, and struggle to generate a large enough customer base just to survive (mainly through promotion and competitive price). This is relevant to the strategy on how to design a service delivery system. For the purpose of ensuring consistency of quality and lowering costs, the de-coupling of back-office activities from the front office has been a predominant strategy in a service delivery system (e.g. Chase, 1978; Chase and Tansik, 1983; Metters and Vargas, 2000). Operationally, this de-coupling strategy divides high- and low-contact activities, segregating/departmentalising and centralizing low-contact activities (Chase, 1978). Within a global context, the decoupling processes may mean that back-office operations are in different countries than the front office. Recent years have also witnessed that the line between front- and back-office are being blurred, with on- and off-line activities becoming more integrated in some situations and more de-coupled in others (Metters and Vargas, 2000). The level of back- and front- office de-coupling can thus take the form of the degree of interaction and customisation. All the above characteristics are fitting well in the conceptualisation of postponement in a supply chain (Yang and Burns, 2003). However, a challenging question arises regarding how postponement should be

defined in the service production and delivery process with the relatively high level of customer participation or interaction in service. Addressing the above question may also provide insight into the impact of customer contact and the existing services linkage in the supply chain and network.

References

- Aviv, Y. and Federgruen, A. (2001), "Design for postponement: a comprehensive characterization of its benefits under unknown demand distributions", *Operations Research*, 49 (4), pp. 578-598.
- Bailey, J. P. and Rabinovich, E. (2005), "Internet book retailing and supply chain management: an analytical study of inventory location speculation and postponement", *Transportation Research Part E: Logistics and Transportation Review*, 41 (3), pp. 159-177.
- 3. Baldwin, C. Y. and Clark, K. B. (1997), "Managing in an age of modularity", *Harvard Business Review*, 75 (5), pp. 84-93.
- 4. Bask, A. H. (2001), "Relationships among TPL providers and members of supply chains a strategic perspective", *Journal of Business & Industrial Marketing*, 16 (6), pp. 470-486.

- Bitner, M. J., Brown, S. W. and Meuter, M. L. (2000). "Technology infusion in service encounters", *Journal of the Academy of Marketing Science*, 28 (1), pp. 138-149.
- Biyalogorsky, E., Carmon, Z., Fruchter, G. E. and Gerstner, E. (1999), "Overselling with opportunistic cancellations", *Marketing Science*, 18 (4), pp. 605-610.
- Blackstone, J. H. (1989), *Capacity Management*, South-Western Publishing, Cincinnati, OH,
- 8. Bowersox, D. J. and Morash, E. A. (1989), "The integration of marketing flows in channels of distribution", *European Journal of Marketing*, 23 (2), 58-67.
- 9. Bucklin, L. P. (1965), "Postponement, speculation and the structure of distribution channels", *Journal of Marketing Research*, 2 (1), 26-31.
- 10. Chase, R.B. (1978), "Where does the customer fit in a service operation?", *Harvard Business Review*, 56 (4), pp.137-42.
- 11. Chase R. B. and Tansik, D. A. (1983), "The customer contact model for organization design", *Management Science*, 29 (9), pp. 1037–1050.
- Choi, T. Y. and Hong, Y. (2002), "Unveiling the structure of supply networks: case studies in Honda, Acura, and DaimlerChrysler", *Journal of Operations Management*, 20 (5), pp. 469-493.
- Daugherty, P. and Pittman, P (1995), "Utilization of time-based strategies: creating distribution flexibility/ responsiveness, *International Journal of Operations and Production Management* 15 (2), pp. 54–60.
- Duclos, L. K., Siha, S. M. and Lummus, R. R. (1995), "JIT in services: a review of current practices and future directions for research", *International Journal of Service Industry Management*, 6 (5), pp.36-52.

- 15. Duray, R., Ward, P. T., Milligan, G. W. and Berry, L. (2000), "Approaches to mass customization: configurations and empirical validation", *Journal of Operations Management*, 18 (6), pp. 605-625.
- Elliott, S. S. and Hughes, B. (2000), "High-Velocity New Product Development", *Proceeding of the 9th International Conference on Management of Technology*, 20-25 February, Miami, USA.
- 17. Eppen, G. (1979), "Effects of centralization on expected costs in a multi-location newsboy problem", *Management Science*, 25 (5), pp. 498-501.
- 18. Feitzinger, E. and Lee, H. L. (1997), "Mass customization at Hewlett-Packard: the power of postponement", *Harvard Business Review*, 75 (1), 116-121.
- Ferdows, K., Lewis, M. A. and Machuca, J. A. D (2004), "Rapid-fire fulfilment", *Harvard Business Review*, 82 (11), pp. 104-110.
- Fisher, M., Hammond, J. H., Obermeyer, W. R. and Raman, A. (1994), "Malting supply meet demand in an uncertain world", *Harvard Business Review*, 72 (3), pp. 83-93.
- 21. Fisher, M. and Raman, A. (1996), "Reducing the cost of demand uncertainty through accurate response to early sales," *Operations Research*, 44 (1), pp. 87-99.
- 22. Fisher, M., and Ittner, C. D (1999), "Impact of product variety on automobile assembly operations," *Management Science*, 45 (6), pp. 771-786.
- 23. Ford, D. (2002), "Solving old problems, learning new things and forgetting most of them: distribution, internationalization and networks", *International Marketing Review*, 19 (3), pp. 225-235.
- Galunic, D. C. and Eisenhardt, K. M. (2001), "Architectural innovation and modular corporals forms", *Academy of Management Journal*, 44 (6), pp. 1229-1249.

- 25. Gattorna, J. L. and Walters, D. W. (1996), *Managing the Supply Chain A Strategic Perspective*, Macmillan, London.
- Gilmore, J. H. and Pine, J. B. (1997), "The Four Faces of Mass Customization", *Harvard Business Review*, 75 (1), pp. 91-101.
- 27. Goldratt, E. (1990), The Theory of Constraints, North River Press, New York.
- 28. Gunasekaran, A., Patel, C. and Tirtiroglu, E. (2001), "Performance measures and metrics in a supply chain environment", *International Journal of Operations and Production Management*, 21 (1/2), pp. 71–87.
- 29. Hagel III, J. (2002), "Leveraged growth: expanding sales without sacrificing profits", *Harvard Business Review*, 80 (10), pp. 68-77.
- 30. Harland, C.M. (1996), "Supply chain management: relationships, chains and networks", *British Journal of Management*, 7, pp. S63-S80.
- 31. Hill, A. V., Collier, D. A., Froehle, C. M., Goodale, J. C., Metters, R. D. and Verma, R. (2002), "Research opportunities in service process design", *Journal of Operations Management*, 20 (2), pp. 189-202.
- Hoekstra, S. and Romme, J. (1992), Integrated Logistics Structures: Developing Customer Oriented Goods Flow, McGraw-Hill, London.
- 33. Huang, M. G., Chang, P. L. and Chou, Y. C. (2005), "A first introduction of customer-driven capacity exchange for job shops", *International Journal of Production Research*, 43 (13), pp. 2743–2758.
- 34. Jin, B. (2004), "Achieving an optimal global versus domestic sourcing balance under demand uncertainty", *International Journal of Operations & Production Management*, 24 (12), pp. 1292-1305.
- 35. Kimes, S. E. (1989), "Yield management: a tool for capacity-considered service firms", *Journal of Operations Management*, 8 (4), pp. 348-363.

- Kotha, S. (1996), "Mass-customization: a strategy for knowledge creation and organizational learning", *International Journal of Technology Management*, 11 (7/8), pp. 846–858.
- Krajewski, L., Wei, J. C. and Tang, L. L. (2005), "Responding to schedule changes in build-to-order supply chains", *Journal of Operations Management*, 23 (5), pp. 452-469.
- Lampel, J. and Mintzberg, H. (1996), "Customizing customization", Sloan Management Review, 38 (1), pp. 21-30.
- Lee, H. L. (1998), "Postponement for mass customization: satisfying customer demands for Tailor-made products. In John Gattorna (ed.), *Strategic Supply Chain Alignment*, Brookfield, VT: Gower, 77-91.
- 40. Liker, J. K. and Choi, T. Y. (2004), "Building deep supplier relationships", *Harvard Business Review*, 82 (12), pp. 104-113.
- Maister, D. (1976), "Centralization of inventories and the 'square root law", *International Journal of Physical Distribution & Materials Management*, 6 (3), pp. 124-134.
- Matthews, H., Hendrickson, C. (2002), "The economic and environmental implications of centralized stock keeping", *Journal of Industrial Ecology*, 6 (2), pp.71-81.
- 43. Metters, R. and Vargas, V. (2000), "A typology of de-coupling strategies in mixed services", *Journal of Operations Management*, 18 (6), pp. 663–682.
- 44. Mintzberg, H. (1979), *The Structuring of Organizations*, Prentice-Hall, Englewood Cliffs.

- 45. Naylor, J. B., Naim, M. M. and Berry, D. (1999), "Leagility: integrating the lean and agile manufacturing paradigms in the total supply chain", *International Journal of Production Economics*, 62 (1-2), pp. 107-118.
- 46. Olhager, J., Rudberg, M. and Wikner, J. (2001), "Long-term capacity management", *International Journal of Production Economics*, Vol. 69 No.2, pp.215-225
- 47. Olhager, J. (2003), "Strategic positioning of the order penetration point", *International Journal of Production Economics*, 85 (3), pp. 319–329.
- 48. Oracle Corporation (2004), "The adaptive supply chain: postponement for profitability", http://productfinder.infworld.com.
- Pagh, J. D. and Cooper, M. C. (1998), "Supply chain postponement and speculation strategies: How to choose the right strategy", *Journal of Business Logistics*, 19 (2), pp. 13-33.
- 50. Peppers, D. and Rogers, M. (1997), Enterprise One to One, Dobleday, New York
- 51. Sako, M., Lamming, R. and Helper, S. R. (1994), "Supplier relations in UK car industry: good news-bad news", *European Journal of Purchasing & Supply Chain Management*, 1, pp. 237–248.
- 52. Sasser, W. E. (1976), "Match supply and demand in service industries", *Harvard Business Review*, 54 (6), pp. 133-140.
- Schilling, M. A. and Steensma, H. K. (2001), "The use of modular organizational forms: An industry-level analysis", *Academy of Management Journal*, 44 (6), pp. 1149-1168.
- Skipworth, H. and Harrison, A. (2004), "Implications of form postponement to manufacturing: a case study", *International Journal of Production Research*, 42 (10), pp. 2063-2081.

- 55. Shingo, S. (1981), *Study of Toyota Production System from Industrial Engineering Viewpoint*, Japan Management Association, Tokyo.
- 56. Tu, Q., Vonderembse, M., Ragu-Nathan, T. and Ragu-Nathan, B. (2004), "Measuring modularity-based manufacturing practices and their impact on mass customization capability: a customer-driven perspective", *Decision Sciences*, 35 (2), pp. 147-168.
- 57. Umble, M. M., and Srikanth, M. L., (1990), *Synchronous Manufacturing: Principles for World Class Excellence* (Cincinatti, OH: South-Western).
- 58. Van Hoek, R.I. and Weken, H.A.M. (1998), "The impact of modular production on the dynamics of supply chains", *International Journal of Logistics Management*, 9 (2), pp. 35-50.
- 59. Van Hoek, R. I., Commandeur, H. R. and Vos, B. (1998), "Reconfiguring logistics systems through postponement", *Journal of Business Logistics*, 19 (1), 33-54.
- 60. Van Hoek, R. I., Peelen, E. and Commandeur, H. R. (1999), "Achieving mass customization through postponement: a study of international changes", *Journal of Market Focused Management*, 3 (3/4), pp. 353–368.
- 61. Van Hoek, R. I. (2001), "The discovery of postponement: a literature review and directions for research", *Journal of Operations Management*, 19 (2), 161-184.
- 62. Van Mieghem, J. A. and Dada, M., (1999) "Price versus production postponement: capacity and competition", *Management Science*, 45 (12), pp. 1631-1649.
- 63. Waller, M. A., Dabholkar, P. A. and Gentry, J. J. (2000), "Postponement, product customization, and market-oriented supply chain management ", *Journal of Business Logistics*, 21 (2), pp, 133-160.

3
4
4 5 7 8 9 10 11
5
0
1
8
9
10
11
12
13
14
15
16
10
17
18
12 13 14 15 16 17 18 19
20
20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
22
23
24
25
26
20
21
28
29
30
31
32
33
34
35
36
37
20
20
39
40
41
42
43
44
45
46
47
48
49
49 50
51
52
53
54
55
56
57
58
59

- 64. Ward, A., Liker, J.K., Cristiano, J.J. and Sobek, D. II (1995), "The second Toyota paradox: how delaying decisions can make better cars faster", *Sloan Management Review*, 36 (3), pp.43-71.
- 65. Wikner, J. and Rudberg, M. (2005), "Integrating production and engineering perspectives on the customer order decoupling point", *International Journal of Operations & Production Management*, 25 (7), pp. 623-641.
- 66. Yang, B. and Burns, N. D. (2003), "Implications of postponement for the supply chain", *International Journal of Production Research*, 41 (9), 2075-2090.
- 67. Yang, B., Burns, N.D. and Backhouse, C. J. (2004a), "Postponement: a review and an integrated framework", *International Journal of Operations & Production Management*, Vol. 24. No. 5, pp. 468-487.
- Yang, B., Burns, N. D. and Backhouse, C. J. (2004b), "Management of uncertainty through postponement", *International Journal of Production Research*, 42 (6), pp. 1049-1064.
- Yang, B., Burns, N. D. and Backhouse, C. J. (2005), "An empirical investigation into the barriers to postponement", *International Journal of Production Research*, Vol. 43 (5), pp. 991-1005.
- 70. Zinn, W. (1990), "Developing heuristics to estimate the impact of postponement on safety stock", *International Journal of Logistics Management*, 1 (2), pp. 11-16.