Postponement Application in the Fast Fashion Supply Chain: A Review

Md Afzalul Aftab¹, Qin Yuanjian¹ & Nadia Kabir²

¹ School of Management, Wuhan University of Technology, Wuhan, P. R. China
² Nanjing, Jiangsu Province, China

Correspondence: Md Afzalul Aftab, School of Management, Wuhan University of Technology, Wuhan, China.
E-mail: afzalulaftab@yahoo.co.uk

Received: March 18, 2017            Accepted: May 19, 2017            Online Published: June 7, 2017
doi:10.5539/ijbm.v12n7p115          URL: https://doi.org/10.5539/ijbm.v12n7p115

Abstract

The successful implementation of push-pull supply chain management strategy has an important role in improving the competitiveness of an organization. The objective of a push-pull strategy is to minimize the holding of inventory level in finished form and rather produce finished goods from semi-finished inventory only upon receiving final order. One of the vital building blocks of push-pull supply chain strategy is postponement. The main objective of this review paper is to discuss the concept of postponement and its sub-categories such as product postponement and process postponement and their benefits. Then it is investigated how two prominent fast fashion retailers who are also categorized as original brand manufacturers in the apparel value chain apply the two variants of process postponement e.g. process standardization and process re-sequencing in their manufacturing operations to activate push pull supply chain strategy. The push-pull supply chain strategy in turn helps to reduce their order-to-delivery lead time to stores, reduce inventory holding level and minimize both physical costs and market mediation costs. The paper ends with concluding remarks. A framework is developed to illustrate the push-pull supply mechanism. This paper is a useful resource for practitioners in apparel supply chain willing to remove inefficiencies, costs and risks in their operations.

Keywords: postponement, fast fashion, apparel industry, supply chain

1. Introduction

In the age of increased globalization and shortened product life cycles, firms in order to meet the diverse needs of global customers are faced with increasingly sophisticated customer demand such as high degree of product variety, requirement for product customization and the requirement to deliver customized products very fast (Swaminathan & Lee, 2003). This puts firms in huge pressure because in order to meet customer’s increased demand for product variety they need to develop, produce and deliver products within short time duration. In such an environment the prediction accuracy for demand, supply and production falls to lower levels which in turn put undesirable effect on firm-level performance. For instance lower accuracy in demand predictability mandates a producer firm to raise the inventory level stock for all types of finished goods to meet required customer service level. This subsequently, can cause larger end-of-life write-offs due to high risk of obsolescence. Furthermore, demand for increased product variety within short delivery lead-time reduces the ability of companies to take advantage of scale economies, therefore making component procurement and product manufacturing difficult to be done in a cost-effective manner. All these factors, coupled with increased competition, put a negative impact on profit margins (Venkatesh & Swaminathan, 2004). Additionally, McCutcheon, Raturi, & Meredith, (1994) presented through a case study that a firm may lose market share to competitor if the firm is unable to meet the demand for product variety in a cost-effective manner while its competitor can. Apparel industry is specifically prevalent with the stated problems of short product life cycle, high product variety, low demand forecast accuracy and finally significant market-mediation cost of holding below or above required inventory levels (e.g. inventory holding cost, markdown) all of which contributes to considerable inefficiency in the supply chain. According to KSA (1997), in the apparel supply chain pipeline only 11 weeks out of the 66 week- lead-time were required on the actual value-adding processes, and the remaining weeks were wasted in non-value adding activities such as work-in-process and finished inventories at various stages in the system. The inefficiency resulted in stocking high inventory stocks of unwanted items and
stock-out for the fast-moving high demand items which resulted in estimated losses of $25 billion. The purpose of this paper is to investigate how some highly successful firms in the apparel industry have managed to bypass the inefficiencies and make significant gains by using principles of speculation and postponement mix in their operations. Postponement implies delaying final manufacturing until customer orders are received and thus helps to reduce market risks while speculation implies keeping semi-finished inventory and performing final manufacturing upon receipt to compress order to delivery lead time. The unit of analysis in this paper is limited to lead firms such as fast fashion retailers who are also categorized as original brand manufacturers in the buyer driven clothing value chain (Gereffi & Frederick, 2010). The findings may be insightful to the top management personnel of apparel retailers and manufacturers in the apparel industry. The paper has considerable practical significance in that, it will suggest a framework that models the strategies of how fast fashion retailers and branded manufacturers are using postponement strategies such as process resequencing and process standardization in their production operations for basic apparel products and fashion apparel products respectively to eliminate many of the risks and uncertainties associated with product demand. The adopted method for the paper is literature review with an intention to document and model how successful fast fashion retailers have adopted their operations practices to respond to the fast changing nature of fashion apparel industry. The following database namely Emerald, Google Scholar and Web of Science were searched using the keywords fast fashion, postponement, and supply chain starting from 1984 to March 2016. The recent and old literatures were systematically searched and the most relevant literatures applying to fast fashion apparel businesses were chosen. However this review does not consider unpublished papers or papers in non-academic journals.

The paper is organized as follows: Section 2 discusses the concept of postponement and its benefits. The application of postponement in fashion apparel industry is discussed in Section 3. Section 4 presents the concept of push-pull supply chain. A framework is developed to demonstrate how fast fashion retailers are using postponement to combine push-pull strategy in their production and supply of basic products and fashion products respectively. Finally, conclusions are presented in Section 5.

2. Postponement—a Solution to Demand Uncertainty

Yang & Burns, (2003) stated that supply chains have to be highly responsive to constantly changing market conditions. Alderson, (1950) pioneered the visionary term postponement and believed is one of the most effective strategic mechanisms to manage the risks associated with product variety and uncertain sales and contributes to improve the efficiency of a marketing system. Lee, Billington, & Carter, (1993), Venkatesh & Swaminathan, (2004) and Yang, Burns, & Backhouse, (2004) stated that by postponing/delaying final processing or manufacturing, the commitment of semi-finished inventories to a differentiated end product in terms of form and identity could be postponed to the latest possible point in time/marketing flow until a customer order containing precise order requirements is received. This solves the problem of risk associated with product variety and sales uncertainty and allows filling demand for product variety through mass customization. Engineer to order is the highest level of postponement where the entire design-make-ship cycle is started only when a customer demand is received. However such a postponement strategy is not feasible in many instances because the market lead time may be shorter than the lead time required under this postponement strategy (Yang et al., 2004). Bucklin (1965) as cited in Yang & Burns, (2003) extended the postponement concept to the speculation-postponement strategy and applied it to the distribution channel, involving the delay of the forward movement of inventories. He mainly examined where, when and who should hold inventory in order to reduce risk and cost. He stated that postponement strategy needs to be combined with speculation strategy since total reliance on postponement will lead to long lead times in production and distribution. Speculation holds that changes in form and the movement of goods should be made at the earliest possible time. Speculation basically facilitates capturing scale economies, but it may lead to substantial inventory buildup. The concept of speculation relies on forecasts of future demand and therefore might be considered as a push-approach, while postponement enables the customer’s needs or requirements to affect product design and production process and thus might be described as a pull approach (Shapiro, 1984).

2.1 Postponement: Enablers, Costs, and Benefits

Swaminathan and Lee, (2003) suggested that innovative companies have exploited the power of product and process design, by integrating design with their supply chain operations to gain control of product variety proliferation. They suggested that 80% of the manufacturing cost of the product is determined by the design of the product or the process in which the product is to be manufactured which implies that design is key factor determining manufacturing cost. Feitzinger & Lee, (1997) and Zinn, (1990) suggested that for postponement application, commonality/modularity in the design of product and processes is vital which paves way for
creating postponable points. At these points, companies have to examine product/production features and analyze the costs and benefits of postponement. Once postponable points have been selected, companies can begin to apply postponement in those points by delaying the configuration/ process/ manufacture to a final end product, thereby increasing the flexibility to handle the changing demand for the multiple products (Swaminathan & Lee, 2003; Yang et al., 2004). Using common components and processes from initial stages implies that the products undertake their unique attributes as late as possible (Lee, 1996). For postponement application to be cost effective, easy and successful a firm must be able to tailor the design of a product or the design of process or both according to the market requirements (Swaminathan & Lee, 2003).

2.1. Product Enablers

2.1.1 Product Enablers

Postponement depends to a great extent on how easily one can store inventory in semi-finished forms. This in turn depends on two factors engineered in the product structure namely parts commonality and product modularity (Venkatesh & Swaminathan, 2004). Parts commonality is achieved through standardizing some key components that are shared by multiple products in the product line thus making it easier to create semi-finished products that could cater to multiple end products. Product modularity on the other hand refers to a product architecture where products are created by assembling a set of modules or components. Alternative product variants are created by different combination of modules or components, where each module can assume different flavors and, therefore, alternative combinations can generate a wide variety of end products (Lee et al., 1993; Swaminathan, 2001). Lee et al., (1993) identifies product modularity as a key enabler of postponement. A simple way to utilize product modularity for postponement is to split a product module into two separate modules, so that one module is mutual across products and the other acts to differentiate between products. In some ways the two techniques—parts commonality and product modularity—are alternate ways to effectively handle product variety.

2.1.2 Process Enablers

Postponement can also be achieved through two types of process changes namely process standardization and process re-sequencing (Venkatesh & Swaminathan, 2004). Process standardization encompasses standardizing process steps related with the different product variants in a product line so that all the product variants in the product line (or a subset of it) pass through the same process step. Typically these are the initial steps in the manufacturing process so that products are not differentiated at these steps, and distinct personalities of the products are added at a later stage. Loosely, one can view process standardization as comparable to parts commonality, except now we are dealing with processes (Venkatesh & Swaminathan, 2004). After creating process standardization the next logical step is to bring the process standardization step earlier in a process sequence through process re-sequencing so that the initial process steps are mutual across different product variants. The outcome is that, more common components are added in the beginning of the process and the components or features that create product differentiation through attaining distinct functionalities and characteristics are added later until market demand conditions become known (Venkatesh & Swaminathan, 2004). In order for process standardization and process re-sequencing to be applicable the structure of a certain process must be modular. If a process can be divided into separate sub-steps so that these sub-steps can be performed in either parallel or in different sequence then it is classified as a modular process. In addition to process modularity, the viability of process re-sequencing depends on common or standard components in the product line (Swaminathan & Lee, 2003).

2.2 Postponement Benefits

Constant end market volatility in terms of volume and variety in the end market place creates risk of demand uncertainty for a product. Postponement has been extensively accepted as an effective and powerful operational strategy for improving the trade-off between cost and customer service in the face of increasing product variety, the need for quick response to customers' needs, and a shortening of the product life cycle, all of which increase the complexity of demand forecasting and planning and ultimately contributes to demand uncertainty (Venkatesh & Swaminathan, 2004; Yang & Burns, 2003). Postponement can effectively avoid the costs related to the risk of demand uncertainty a product is faced with (Swaminathan & Lee, 2003). Costs affected by postponement can be broadly classified into costs associated with the amount and nature of inventory stocked (i.e. stocking of semi-finished inventory and finished product inventory) and logistic cost (Rockhold, Lee, & Hall, 1998).

2.2.1 Inventory Perspective

*First*, in postponement application, component level inventory requirement is an aggregation of demand for all finished products that use this component. As such it can be forecasted with accuracy and thus can be stocked through purchasing/producing in advance ultimately eliminating uncertainty in component level demand.
In this section we investigate how the fast fashion retailers who are also brand manufacturers (Gereffi &
3. Postponement Application in Fashion Industry
Cooper, 1998; R. van Hoek, 1997; Yang et al., 2004; Zinn & Bowersox, 1988).
risks all of which are associated with product variety and uncertain sales in a dynamic environment (Pagh &
sequence: (1) dyeing; and (2) knitting. In the first stage, raw materials of white yarns were dyed and converted
Dapiran, (1992) and Lee & Tang, (1997) showed that, Benetton, a major apparel manufacturer used to
3.1 Process Standardization and Process Resequencing Application for Basic Product Category at Benetton
Furthermore the stocked semi-finished components facilitate manufacture/build to order in short lead time which
reduces the need for the level of safety stock of finished goods inventory. Swaminathan & Tayur, (1998) pointed
that the shortened lead-time achieved through postponement and ultimately faced by the customer is an indicator
of improved service level provided to customers. Whang and Lee, (1998) suggested that postponement
effectively delays the point of product customization and thus shortens the forecast horizon for the demand of
end products. The demand predictability of end product inventory is enhanced since short-term forecast tends to
be more accurate than long-term forecast. Therefore it results in additional decrease in finished product
inventory holding level which positively affects the profit and loss statement since costs such as devaluation,
obsolescence, and storage associated with holding finished product inventory are eliminated.
2.2.2 Logistic Perspective
Secondly, concerning the logistics area in the material flow perspective it is viable to delay a product's variety/
weight/ volume/ value - added step to save on inventory carrying and holding, stock-out and obsolescence costs.
Postponement typically reduces the total value of the items held in inventory, since more generic and less
value-added products are held (Johnson & Anderson, 2000). This guarantees to gain leverage in terms of greater
flexibility, efficient asset utilization, effective control of costs, manage service performance and minimizing the
risks all of which are associated with product variety and uncertain sales in a dynamic environment (Pagh &
Cooper, 1998; R. van Hoek, 1997; Yang et al., 2004; Zinn & Bowersox, 1988).
3. Postponement Application in Fashion Industry
In this section we investigate how the fast fashion retailers who are also brand manufacturers (Gereffi &
Frederick, 2010) apply postponement through process standardization and process re-sequencing in their
manufacturing system. Christopher & Peck, (1997) suggested that in the fashion market, factors such as short
product life cycle, high impulse purchasing attitudes and high volatility of demand pattern all lead to market
uncertainty. He suggested that shorter lead times can effectively counter the risks associated with market
uncertainty since shorter lead time implicates that the forecasting horizon will be shorter and consequently the
risk of error will be lower. The first case investigated refers to Benetton-an Italian fashion company and explains
how they initially applied process standardization and process resequencing to sweater production which is a
basic product category. This concept can of course be extended to all basic products category in the apparel
product line. The simple basic style apparel products contain low fashion content, come in few classic colors,
and have product life cycle of two to three years with little seasonal variation that sell all year round e.g. men's
and children's merchandise such as sweater, dress shirts. Basic style apparel products have a stable demand
pattern and can be forecasted with higher accuracy (Doeringer & Crean, 2006; Ghemawat & Nueno, 2006;
Warburton & Stratton, 2002). The second case refers to Inditex-Zara, a Spanish fast fashion retailer and explains
how they apply process standardization to their fashion products category in their apparel product line. Fashion
products are intricately worked garments with higher fashion content and produced in different types of styles,
materials, colors, and silhouettes. Fashion products usually have a short life-cycle of 1 month or less because the
number of seasons per year has increased in numbers thus shortening time span of each season. Therefore
fashion products get only one delivery for replenishment purpose to the stores (Doeringer & Crean, 2006).
Additionally demand for fashion product is highly unpredictable suggesting lower forecast accuracy (Abernathy,
Volpe, & Weil, 2006).
3.1 Process Standardization and Process Resequencing Application for Basic Product Category at Benetton
Dapiran, (1992) and Lee & Tang, (1997) showed that, Benetton, a major apparel manufacturer used to
manufacture sweater using the traditional process which consisted of two major stages in the following
sequence: (1) dyeing; and (2) knitting. In the first stage, raw materials of white yarns were dyed and converted
into yarns of different colors. The dyeing operation was a batch operation in which a large quantity of white
yarns could be submerged into different pools of dyes. In the second stage the dyed/colored yarns were knitted
into diverse finished garment products (different styles and sizes). The garments were stored in the form of
finished goods to be distributed to the retailers. The production process in this sequence is a perfect example of
“make-to-forecast” which is illustrated by the positioning of the customer order decoupling point (CODP1) at the
beginning of the first stage in figure 1.
However several problems arise in this sequence. First Dapiran, (1992) suggested that most of the demand
variability was due to the uncertainty of the customers’ preference of colors in a particular season. The likely
result of this approach is that invariably the desired colors will be out of stock while there are excess inventories
of the unpopular colors. The typical result is costly end of season mark-down and lost sales opportunities. The
lost sales opportunity of popular color products can be blamed upon the sequence of “dye-first knit-later” since
Lee & Tang, (1997) suggested that the processing time per sweater required by the knitting operation is much longer than that required by the dyeing operation. This fact increases the lead-time for replenishment of popular color product which may be greater than market lead-time. Secondly following the same logic that the knitting process is slower compared to the dyeing process, Lee & Tang, (1997) pointed that if the dye-first knit-later sequence is to be followed, it necessitated holding high levels of inventory of finished garments to meet customer service expectations. This not only adds up to the cost of holding more inventory but exposes to the higher risk of unsold inventory of the unpopular colored products.

Harvard Business School Note (1991), Dapiran, (1992), and Lee & Tang, (1998) reported ways in which Benetton addressed the problem of product proliferation. Luciano Benetton, the chairman of Benetton, was credited with his innovative re-engineering of the supply chain by reversing the “dyeing” and “Knitting” stages and enabling postponed dyeing. Lee & Tang, (1997, 1998) consider knitting and dyeing tasks for garments as a two-stage system where at each stage a distinct feature is introduced into the product. Hence in the first stage, yarns that have been pre-treated in a strong chemical solution to increase their responsiveness to dye i.e. bleached yarns are knitted into different styles and sizes of “greige” (uncolored) knit garments, thus allowing stocking semi-finished inventory waiting to be dyed in different colors based upon market response. In the second stage, the semi-finished inventory is dyed into the different colored end products when the season’s fashion preferences become available through EDI. The process is illustrated in Figure 2.

The re-sequencing of the production process to “knit-first dye-later” for the basic product category has several insights. First, the demand for garments falling in the basic product category is relatively more stable and as such the aggregated demands of the product consisting of different sizes can be forecasted with higher accuracy. This is possible because the company can easily collect data regarding the more common body measurement sizes from the demographic data of a specific targeted market in a region. The attribute that contains uncertainty in a basic product is the color it assumes as an end product.

![Figure 1. Dye-first Knit-later sequence](image-url)
Therefore by using the sequence of “knit-first dye-later”, knitting becomes a standardized process and is common to all the product variants in terms of size and style for a particular basic product. It therefore reduces the production variability of the first operation, since essentially a single product is being knitted at that stage (instead of dyeing multiple colored yarn at the first stage)—(see figure 2). Secondly, the re-sequenced production process is a perfect example of “make-to-order” which is illustrated by the re-positioning of the customer order decoupling point (CODP2) at the beginning of the second stage in figure 2. Till this point product differentiation is postponed and only after this point the semi-finished inventory is differentiated by dyeing into different colors after observing the sales data through EDI and having a better idea of popular color of season (Lee & Tang, 1998). Lee & Tang, (1997) stated that the value added during the dyeing stage is higher than that of the knitting stage. As such through postponing or delaying dyeing process, value is added in the supply chain as late as is consistent with meeting customer needs. Third, Lee & Tang, (1997) further stated that the processing time per sweater required by the dyeing operation is much shorter than that required by the knitting operation, because the dyeing operation is a batch process. As such the new production sequence of “knit-first dye-later” significantly reduces the product replenishment lead-time of popular colored garments compared to market lead-time since products have already been knitted in advance and waiting to be dyed. As such Benetton can observe the market trend, identify the popular color garment and replenish accordingly which solves the problem of lost sales opportunities and costly end of season product mark-down. Fourth, due to reduced lead-time for product replenishment of popular colors Lee & Tang, (1997) pointed that it relieves the necessity of accruing costs for holding high levels of inventory of finished garments to meet customer service expectations. Finally, basic products are more price-sensitive than time sensitive (Ghemawat & Nueno, 2006). Therefore in the total cost buildup for basic products, physical costs (e.g. labor cost) are more dominant than market mediation costs (e.g. mark downs due to unsold inventory or lost sales due to stock-outs) which implies that minimizing physical costs should be the focus (Fisher, 1997). A natural choice of production venue for basic products are distant countries where labor wage rates are lower but order to delivery lead times are significantly longer compared to local/proximate sourcing (Ferdows, Lewis, & Machuca, 2004). Also to reduce physical costs, production of basic products can be done in larger batches to achieve scale economies (Sen, 2008). Using the production sequence of “knit-first dye-later”, Benetton has in addition to producing basic products in their own manufacturing facilities, the choice to source the knitted garments in undyed form from cut, make, and trim (CMT) subcontractors located in low cost countries to reduce physical cost and keep in the semi-finished product in stock. After knowing the popular color choice of season, the basic garments can be dyed in own manufacturing facilities. However, Lee & Tang, (1998) argued that the new sequence of “knit-first dye-later” improves operational performance only when the total aggregate demand of garment is predictable with higher accuracy and variability in the style of garment is low although the end product may come in few classic colors and multiple sizes. Such products already mentioned earlier are basic apparel products and fall into functional product category (Fisher, 1997). Lee & Tang,
3.2 Process Standardization Application for Fashion Product Category at Inditex-Zara

Fisher, (1997) categorized high fashion product in the innovative product category because the demand is highly unpredictable and the product category is more time sensitive but less price sensitive. Therefore, Fisher, (1997) suggested to minimize market mediation cost (e.g. mark downs due to unsold inventory and dissatisfied customers/lost sales opportunity due to stock-outs) which dominate over physical cost (e.g. labor/material/inventory storage/transportation cost). Market mediation cost can be minimized by producing in own manufacturing facilities and subcontracting labor intensive process to local cut, make, trim (CMT) subcontractors or sourcing from original design manufacturers/ODM suppliers in geographically proximate countries (Gereffi & Memedovic, 2003; Tokatli, 2008). Such sourcing strategy may be more expensive given fast fashion retailers are both located in and also neighboring to countries with high labor wage rates compared to distant countries with low wage rates. However the attained capabilities of rapid responsiveness to demand changes in terms of volume/ variety and frequent product replenishment option for the fast fashion retailer help attain fast inventory turn around and minimize market mediation costs which in turn counter balances higher manufacturing labor costs (Abernathy, Dunlop, Hammond, & Weil, 1999).

Inditex-Zara, the leading fast fashion retailer and also a brand manufacturer (Gereffi & Frederick, 2010), utilizes process standardization at two very important steps that significantly reduces product development and order to delivery lead time.

First Zara utilizes process standardization by storing undyed fabric inventory in the warehouse, which is later adapted into different colors by dyeing in own manufacturing facilities after receiving precise customer order specifications (BusinessWire, 2002). Fabric supply is an essential part of the overall supply process, with significantly longer lead times than garment production cycles. Instead of forecasting at the detailed mix level of garment products which is likely to be inaccurate, it is easier to accurately forecast the overall demands (the number of aggregate items) (MacCarthy & Jayarathne, 2009). Imprecise volume estimation results in either a network that cannot supply the required volumes because of capacity limitations or result in costly unutilized spare capacity (Martin Christopher, Lowson, & Peck, 2004; Tyler, Heelley, & Bhamra, 2006). Zara uses the information of target sales volumes to purchase the required volume of fabrics in advance enabling the garment production system to be decoupled from the longer lead time of fabric production system. Not only responsiveness is enhanced significantly, the risk is also reduced significantly since numerous garment styles may often be produced from a specific fabric type (MacCarthy & Jayarathne, 2009; Yang & Burns, 2003). MacCarthy & Jayarathne, (2009) noted that a key issue that needs to be considered is the coloring process and pointed that garments may be:

1. Manufactured from fabric that does not need coloring;
2. Manufactured from fabric which is subject to coloring after fabric production; or
3. Finished garments may be colored once manufactured.

Zara uses the second process regarding the fabric coloring. Zara stores 50% of its fabrics in a ‘grey’ undyed state so that it can react faster to mid-season color changes and wait until orders are confirmed before transferring fabric for dyeing process (Ferdows et al., 2004). This strategy Zara uses is known as “process standardization” to gain more speed and flexibility (Yang & Burns, 2003). Much of this volume is funneled through Comditel, a 100%-owned subsidiary of Inditex (Zara’s parent company group) that deals with more than 200 external suppliers of fabric and other raw materials. Comditel manages the dyeing, patterning, and finishing of gray fabric for all of Inditex’s chains, and not just for Zara. To expedite quick changes in printing and dyeing, Zara also work closely with Fibracolor (a dye stuff producer part owned by Inditex - Zara purchase 20% of its output). The rest of the fabrics come from a range of 260 other suppliers, none account for more than 4% of Zara’s total production in order to minimize any dependency on single suppliers and encourage maximum responsiveness from them (Ghemawat & Nueno, 2006). The dyed fabrics are cut via computer aided cutting machines in Zara’s own manufacturing facilities and sent to a network of 500 local assembly subcontractors performing the labor intensive cut, make and trim (CMT) process. Subcontractors then transport back the sewn items to a Zara facility, where each piece is inspected during ironing (by machine and by hand), given a machine-readable tag and packed. The finished products are then sent to the distribution center from where products get shipped to stores (Ferdows, Lewis, & Machuca, 2003). Zara also sources from ODM suppliers located in geographically
proximate Eastern European countries, especially Turkey, Morocco and Bulgaria that helps to compress order-to-
delivery lead times. As such Zara is able to produce/source fashion products in small lots and deliver to stores to
understand market response. If the product is a failure, the financial loss is minimized as the merchandise in
stores are in small quantities. If the product turns out to be successful it can be replenished in very short time
(Tokatli, 2008).

Secondly in order to generate large variety of products and offer quick response to customers, Zara also utilizes
process standardization in the design phase of its products through developing standardized design modules also
known as vanilla box designs (Fraiman & Singh, 2002; Pich, Hay den, & Harle, 2002). At the beginning of each
selling season, the designers create a library of design modules that serve as platforms also known as vanilla
boxes for the models that will be eventually launched. The concept design modules are stored in image format in
computers and do not represent any physical inventory (Venkatesh & Swaminathan, 2004). Designers at Zara
walk the streets, go to discos, and visit local events in order to get a feel for the latest fashion trends. After
cautiously spotting the latest trends, the designers give adaptation or customization to the library-held vanilla
box designs, and create nearly 5-8 new designs every day (Venkatesh & Swaminathan, 2004). If the vanilla
designs were not obtainable to the designers, the design process could take several days. Even though there is no
physical inventory, the design by vanilla box approach helps rationalize the complexity of product variety that is
inherent in the fashion apparel industry and help manage the short product life cycle. In total, about 12,000 new
products and designs are created every year (Ghemawat & Nueno, 2006). Zara is an example of process
standardization adapted at the design-manufacturing interface, so that the time of differentiation/production can
be postponed (Venkatesh & Swaminathan, 2004). In the fashion apparel industry it is common to commit to
production two years ahead of the actual sales season, instigating considerable losses due to stock-outs and
inventory obsolescence (Fisher & Raman, 1996). This is especially upsetting in an industry with short product
life cycles of one season lasting less than 1 month. However, Zara is able to introduce new products at a rapid
rate; in fact 70% of the products change every two weeks in a typical retail outlet. Zara is able to introduce new
products or replenish existing successful products belonging to the high fashion category in small quantities on a
frequent basis (at least twice a week) thus maintaining freshness and exclusivity to their products (Abernathy et
al., 2006). The attained operational flexibility and speed to produce a variety of styles in smaller quantities in
shorter lead time allows Zara immediate response to emerging fashion trends and send on-time delivery of
fashion products to stores (Fisher, 1997).

In the following section we discuss the various concepts such as supply chain time line, the product/information
flow that lies within the supply chain time line and the associated material/information decoupling points. Then
we discuss how Benetton and Zara’s process standardization and process re-sequencing strategies to achieve
postponement application conforms to the supply chain fundamentals that are prescribed to achieve lean retailing
benefits (i.e. inventory minimization).

4. Push–pull Supply Chain and Material/Information Decoupling Points

A supply chain is a network of facilities that performs the functions of procuring material, transforming material
to intermediate and finished products, and distributing the finished products (Lee & Billington, 1994; 
as the time that elapses between the procurement of raw material, that is, the beginning of the time line, and the
delivery of an order to the customer, that is, the end of the time line (see figure 3). Mason-Jones & Towill, (1999)
suggested that supply chains have at least two distinct flow pipelines:

1. The product transfer pipeline from raw materials to end customer.
2. The order information transfer pipeline, from point-of-sale through to raw material supplier.

Information regarding demand of end customer in the marketplace and resultant order information flows
upstream. Using this information production is activated and the products flow downstream. Therefore speed and
reliability of order data transference is critical to enabling good supply chain dynamics.

4.1 Material Decoupling Point

The material flow decoupling point is a specific point along the product transfer pipeline where the firm switches
from managing the supply chain using a push strategy to managing it using a pull strategy. Yeh & Lee, (2016)
termed this point as push–pull boundary. Swaminathan and Lee (2003) said that the material flow decoupling
point can be used to establish postponement strategy and thus termed as point of differentiation. Mason-Jones &
Towill, (1999) pointed that till this point semi-finished inventory/generic products are produced to speculation
i.e. (using push strategy) to be stocked at this point. At this point in the product axis, customer’s order penetrates
and using this information the semi-finished inventory is configured to finished products i.e. (builds to order using pull strategy). Because the inventory is generic, its flexibility is greater, meaning that the same components, modules or platforms can be embodied in a variety of end products (Yang et al., 2004). Modules with standardized interfaces could be combined into customer-specific finished products only upon receiving customer order that specify exact product information regarding mix and match. This permits a manufacturer to attain the marketing benefits of customization while reaping the cost benefits of standardized production of modules with standard interfaces in anticipation of future customer orders (Lampel & Mintzberg, 1996). This strategy is named mass customization, where the key is postponing the task of differentiating a product for a specific customer until the latest possible point. Therefore we find that postponement is an enabler of mass customization that ultimately enhances a company’s flexibility by enabling customization of products and services (Feitzinger & Lee, 1997; Kotha, 1995; Lampel & Mintzberg, 1996; Van Hoek, 1999). Ulrich, (1995), based on his experience with the bicycle industry indicated that the proper place to locate the point of differentiation is upstream of the point where most of product variety is introduced into the supply chain. Yang et al., (2004) suggested that for postponement and mass-customization strategies to be effective, product design needs to separate the most variable portion of the functionality from other functions in order to be added last. Also, features with long lead times need to be standardized and performed earlier in the chain and features with short lead times that add differentiation features to a product can be performed at a late time point closer to the customer. Feitzinger & Lee, (1997) added that since in mass customization, manufacturing is being postponed which suggests that customer orders will be fulfilled through production rather than through stockholding finished products, a reliable supplier network that can supply parts (e.g. generic base product, standard and modular components) and services (e.g. meeting supply quality and delivery promise) is crucial for postponement applications. This reduces supplier-induced uncertainty towards both manufacturing and logistics postponement and improves flow of material across business functions (Feitzinger & Lee, 1997; Yang & Burns, 2003). Additionally, postponement centers on delaying the final customization or forward positioning of finished goods (i.e. logistics postponement) until customer requirement or the final market destination is known. This often necessitates acceleration of production and premium movement (e.g. changing the mode of transportation from shipping to air freight) of specific components or finished products through the system which can result in significant increase in transport cost (Christopher, 1998).

4.2 The Information Decoupling Point

Braithwaite, (1993) suggested that to fully realize performance improvement potential, supply chains need to develop integrated strategies for both information and material flow pipelines. Information flow is the order transmission from point of sale through to raw materials suppliers in the supply chain (Yang & Burns, 2003). Braithwaite, (1993) discussed the example of a company striving to reduce manufacturing cycle times by just one day while not attempting to tackle 2-3 week ordering delays which worsened the customer service levels. Forrester, (1960) found that market sales information particularly suffers from delay and distortion as it moves upstream through the supply chain. The result is production profiles at the factory, which bear little resemblance to the end customer’s buying behavior. Thus creating problems in the material flow pipeline. Given such disguising of what is happening at the marketplace, the need for transparent information flow is of utmost importance. Mason-Jones & Towill, (1999) stated that information flow pipeline has its own significant decoupling point namely information decoupling point. They defined it as the point in the information pipeline where the information of market place order data penetrates without modification and meets with forecast driven information flows. Mason-Jones & Towill, (1999) stated that all supply chains have undistorted order information obtainable at the marketplace interface dealing with the end customer i.e. retail store position. If the order decisions within each player in the supply chain could be enriched with the unbiased, undistorted and rich information of market sales data available downstream via moving the information decoupling point upstream, the information distortion and wave form promulgation often exhibited in the real world could be greatly reduced. This would positively impact the behavior of both ordering and stock level dynamics from an individual business and from a total supply chain perspective level. Mason-Jones & Towill, (1999) suggested that to maximize the strategic potential of these data within the supply chain, the ideal place for information decoupling point is at the furthest point upstream possible enabling the maximum number of members in the chain to be linked directly with the market place and avail the actual market place information for better forecasting and decision making. Lee, (1998) suggested that quick and accurate capture of the customers’ order information in undistorted form available downstream (e.g. retail outlets) and transmitting it upstream (e.g. headquarters) in a timely manner is extremely valuable to postponement application. Information technology such as internet, electronic data interchange (EDI) and point-of-sale system may improve the information flow by reducing the data collection errors and moving data quickly. However, they are not sufficient to eliminate distortions or
identify good information in the information flow (Yang et al., 2004). All the players in the supply chain need an overall supply chain vision (Van Hoek, 1999); share a common target (Hong-Minh, Disney, & Naim, 2000); recognize the need for open communication to enable information sharing; must be willing to act as partners in the supply chain and finally need a policy on the strategic placement of the information decoupling point (Mason-Jones & Towill, 1999). Organizational administrative culture must also be supportive of providing undistorted information to their respective suppliers and customers (Towill, 1996; Van Hoek, 1999). The resulting outcome will be that enhanced information flow will reduce the uncertainty induced by the bullwhip effect (Lee et al. 1997) that in effect will eliminate control uncertainty which is concerned with how internal decision making affects the ability to transform customer orders into activities in the supply chain, e.g. supply raw material requests (Towill, Childerhouse, & Disney, 2002). Finally reduced uncertainty will facilitate more effective application of postponement (Yang & Burns, 2003). Figure 3 summarizes the perception of the material decoupling point and information decoupling point and their relative positions within the supply chain time line.

Figure 3. A framework illustrating push-pull supply in apparel manufacturing and supply operations
Source: Authors.

4.3 Fast fashion Retailer’s Push-Pull Supply Strategy Illustration

Fast fashion retailers such as Benetton and Inditex-Zara who are also brand manufacturers (Gereffi & Frederick, 2010) utilizes the two decoupling points namely material decoupling point and information decoupling point to apply combination of speculation/postponement strategy to perform order-driven production from semi-finished inventory rather than forecast-driven production. This is enabled by their strategic use of both process standardization and process resequencing for the basic products and process standardization for fashion products in their product line. Both Benetton and Zara have integrated information system in place to transmit and exchange information among various units. For instance the point of sales (POS) data are collected from each retail stores via bar code readers and transmitted via electronic data interchange (EDI) to the headquarters. The head quarter in turn can transmit information to manufacturing facilities, distribution centers and stores all of which in turn can respond in real time. All the business units (e.g. retail stores, head quarter, manufacturing units, suppliers and distribution centers) of modern retailers following lean retailing business model have strong IT integration (Abernathy et al., 1999). The mechanism is illustrated in figure 3. Benetton produces basic products and Inditex-Zara produces fashion products in their own manufacturing facilities. Zara usually out sources basic
apparel items from low wage countries in Asia (Gereffi & Frederick, 2010). The activities related to basic products are marked in brackets with subscript(B); fashion products with subscript(F) and activities related to both basic and fashion products with subscript(B,F). The demonstrator model in figure 3 used by the brand manufacturers/fast fashion retailers, utilizes a four level supply chain constituting a head-quarter, manufacturing facility, distribution center and retail stores. For both the basic products and fashion products we find that greige/white fabric is purchased in advance of production since fabric production adds up to long lead-time (MacCarthy & Jayaratne, 2009). Standardized processes related to basic products such as design are performed at head quarter and cut and knit performed in own manufacturing facilities. All of these steps are done in advance of order and kept as semi-finished inventory (Lee & Tang, 1998). This is possible because basic products’ variable features constitute mostly of colors and sizes and not much of style/shape/fashion content (Yeh & Lee, 2016). Retail stores after learning about popular color of season from sales trend transfer the undistorted order information to the head-quarter. The head-quarter transmits information to the manufacturing facilities where the undyed knitted clothes are dyed according to popular color of season and the operations such as finishing and packaging are performed and sent to distribution center for distribution and delivery to retail stores (Ghemawat & Nueno, 2006). For the fashion products, Zara utilizes process standardization in creating vanilla box designs which means standardized design modules that serve as platforms. Zara’s design team after understanding the latest fashion trend from fashion shows and its own retail stores, gives adaptation or customization to the library-held vanilla box designs (Venkatesh & Swaminathan, 2004). The vanilla box design creation and design adaptation are done in advance of order at the headquarter; capital intensive activities such as cutting and dyeing performed in own manufacturing facilities; labor intensive stages such as knitting, finishing and packaging outsourced to local CMT subcontractors and finally distribution performed through own distribution centers respectively upon receiving order at the material decoupling point. The order information is transferred to the head quarter from the retail stores via EDI (Ferdows et al., 2004). Figure 3 shows that the information decoupling point is positioned at the farthest point upstream i.e. at the head-quarter, thereby allowing all four members of the supply chain i.e. the head-quarter, factory, distribution center and retail stores to have an undistorted view of the consumer buying behavior. All the members in the supply chain is forewarned of changes in sales profile and is able to respond in a dynamic and effective manner to meet the demands while avoiding the added burden of the distortion of end customer demand characterized by the Forrester Effect. In this supply chain, information transparency is in place before product assumes its differentiated form according to market requirement at the material flow decoupling point. This indeed is an illustration of the optimized order information pipeline and may be called the information enrichment model, achieved via an electronic point of sales link (EPOS) to all players downstream of the information decoupling point (Mason-Jones & Towill, 1999). The utilization of both the EDI enabled information decoupling point and material decoupling point approaches has enabled the fast fashion retailers’ supply chain to effectively apply postponement but not at the expense of increased order to delivery lead time.

5. Conclusion

The textile and apparel industry is characterized by short product life cycles, high product variety and increased competition. To stay competitive, the extended organizations have to match supply with demand to reduce obsolescence and opportunity cost. After a literature review, this paper has attempted to address the question of how postponement is being applied in the fast fashion apparel business to reduce both physical costs and market mediation costs and proposes a framework illustrating the mechanism. This paper has reviewed two cases of fast fashion retailers namely Inditex-Zara and Benetton who are also brand manufacturers and found that they have been able to significantly reduce demand uncertainty and thus market mediation costs through application of push-pull supply chain strategy. Push-pull supply strategy is activated through application of “process re-sequencing” and “process standardization” for basic apparel and process standardization for fashion apparel production; strong information infrastructure in place and use of material decoupling point and information decoupling point concepts. As such, product supply to the stores can be closely matched with market demand. The case descriptions would be helpful for the practitioners in the industry in order to learn about the management practices supporting different postponement strategies. Similarly the study presents an opportunity to the researchers to look at postponement application within the apparel and textile industry and to compare the practices with postponement application in other industries. It is proposed that rigorous statistical analysis can be performed upon big samples of fast fashion retailers/brand manufacturers to verify how postponement application affects their operational and financial performance.

Acknowledgements

The authors are most grateful to Professor Qin Yuanjian for his constructive and helpful comments on the paper,
which helped to considerably improve the presentation of the paper.

**References**


Towill, D. R., Childerhouse, P., & Disney, S. M. (2002). Integrating the automotive supply chain: where are we


Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/4.0/).