Responsiveness and Efficiency of Pull-Based and Push-Based Planning Systems in the High-Tech Electronics Industry

Dennis Minnich
McKinsey & Company, Inc.
Königsallee 60C
40027 Düsseldorf, Germany
e-mail dennis.minnich@i-u.de

Frank H. Maier
International University in Germany
Campus I
76646 Bruchsal, Germany
Phone +49 7251 700-320
Fax +49 7251 700-350
e-mail frank.maier@i-u.de

Abstract

Planning systems for supply chain management can be designed in different ways in order to achieve the multiple objectives of supply chain management. Two of these objectives are responsiveness and efficiency of the system. In this paper, a System Dynamics simulation model is used to assess the performance of planning approaches for supply chains in the high-tech electronics industry on the dimensions of responsiveness and efficiency. Supply chains in this industry are subject to a number of challenges that complicate the achievement of these two objectives. The findings suggest that pull-based planning approaches can be used to achieve high efficiency and responsiveness, at the expense of higher fluctuations in capacity utilization upstream in the supply chain.

Keywords: Supply Chain Management, High-Tech Electronics Industry, Planning Approaches, System Dynamics

1. Achieving Efficiency and Responsiveness in Complex Supply Chains

Efficiency and responsiveness are two objectives of supply chain management. Whether or not high performance can be achieved on both dimensions at the same time, however, is questionable. Responsiveness can be defined as the “the ability of the supply chain to respond purposefully and within an appropriate time-scale to customer re-

---

1 Dennis Minnich, MSc, is a consultant at McKinsey & Company, Inc. Previously, he was a Research and Teaching Associate at the Department of Operations Management at the International University in Germany, Bruchsal, and PhD student at the Industrieseminar of the University of Mannheim.

2 Prof. Dr. Frank Maier is Professor of Operations Management and Dean of the School of Business Administration at the International University in Germany, Bruchsal.
quests or changes in the marketplace”.3 In contrast, a supply chain would be considered to be efficient if the focus is on cost reduction and no resources are wasted on non-value added activities.4 Decision makers in supply chains face conflicting priorities and need to weigh and balance various performance objectives, such as delivery performance or supply chain costs, and need to find appropriate ways to improve performance on one or both of these dimensions. Supply chains can be designed in different ways for different types of products in order to address this potential trade-off between responsiveness to customer requirements and efficiency, but they are also subject to external factors and challenges.

Supply chains face a number of challenges that can reduce the responsiveness of the system. Examples include long component lead times, erroneous components, capacity constraints, complex technologies and missing information about actual end customer demand. There are also internal challenges for supply chains, such as mismanagement of the planning process. In particular the high-tech electronics industry is – compared to other industries – confronted with a large number of challenges to supply chain management; the most important ones are summarized in Table 1. The combination of these challenges leads to delays in both the flow of information and the flow of materials and requires long-term planning.5

<table>
<thead>
<tr>
<th>Challenges</th>
<th>Other industries facing these challenges (examples)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short product life cycle and rapid price erosion</td>
<td>Pharma, Retail (fashion apparel)</td>
</tr>
<tr>
<td>Volatile and unpredictable demand, especially for new products</td>
<td>Retail (fashion apparel)</td>
</tr>
<tr>
<td>Stock keeping unit (SKU) proliferation caused by customization requirements</td>
<td>Consumer packaged goods</td>
</tr>
<tr>
<td>Long lead time for key components</td>
<td>Retail (fashion apparel)</td>
</tr>
</tbody>
</table>

Challenges | Other industries facing these challenges (examples)
--- | ---
Demanding customer requirements on lead time and volume flexibility | Automotive
Global low cost country (LCC) based supply chain leading to higher end-to-end complexity | Automotive, Retail (fashion apparel)
Complicated product design involving thousands of components | Automotive
High inventory cost and risk of obsolescence | Automotive

Some of these challenges are not unique and experienced in other industries as well, such as the fashion industry. For example, Christopher et al. as well as Hochmann find that short life cycles, high volatility, seasonality and low predictability, high SKU complexity, long materials lead times and capacity constraints are among the characteristics of the fashion industry.7 Other challenges are present in the consumer packaged goods industry, such as SKU proliferation. Similarly, automotive companies, such as General Motors, faced an environment with national markets and trade barriers until the 1990s. Today, the environment they operate in is perceived to be more global, with fewer trade barriers and higher customer expectations. New entrants to the industry caused and continue to cause a proliferation of model offerings and brands.8 This is similar to the development observed in the high-tech electronics industry. The combination of all of these challenges makes supply chains in the high-tech electronics industry a good example for a complex system, providing a large variety of issues to be analysed.9

In the following, a System Dynamics simulation model is used to identify appropriate ways to design supply chains such that they can cope with these challenges. As Lee notes, “the best supply chains aren’t just fast and cost-effective. They are also agile and adaptable, and they ensure that all their companies’ interests stay aligned”.10 Decision makers in supply chains face conflicting priorities and need to weigh and balance various performance objectives, such as delivery performance or supply chain costs, and need to find appropriate ways to improve performance on one or both of these dimensions. There may be planning approaches that achieve both, increased efficiency and increased responsiveness, at the same time and optimise the balance between the two for

---

8 This was pointed out by G. Richard Wagoner, Chairman and CEO of General Motors, Detroit, during a talk on “Leadership in the Automotive Industry” on 11 October 2006 at the Sloan School of Management. A video recording of the talk is available at http://mitsloan.mit.edu/corporate/dils.php.
different types of products. By developing a System Dynamics simulation model of a multi-tier supply chain system in the high-tech electronics industry, this paper investigates the impact of altering key aspects of the planning activities in the supply chain with the objective of determining the scope for potential improvements in both responsiveness and efficiency of supply chains in this industry. The model developed by the author allows simulation of the dynamics in supply chains as complex as those in the high-tech industry and supports the identification of appropriate supply chain planning approaches for products with different characteristics. While the model serves primarily as a basis for research, research findings are of high practical relevance. The System Dynamics model can support decision makers in managing supply chains according to the goals of responsiveness and efficiency.

2. Planning Approaches in the High-Tech Electronics Industry

2.1. Production Processes and Supply Chain Structure in the High-Tech Electronics Industry

Supply chains for products in the high-tech electronics sector typically stretch from raw materials, such as silicon for wafer manufacturing and crude oil for the plastic used for injection molded cases, through several different component manufacturers who supply the final assembly plant with components. From the final assembly plant, the products pass through companies such as network operators, distributors and/or electronic goods retailers to reach the end customer and ultimate user of the product. There are three main production steps for most high-tech electronics products, such as network routers, mobile phones or MP3 players. These are (1) board printing, also known as printed circuit board assembly (PCBA), (2) final assembly and (3) software flashing, packaging and shipment. The final step of software flashing and packing is typically done in house by OEMs such as Nokia, Ericsson, Apple or SonyEricsson, and a large percentage of the rest of the supply chain is often outsourced to contract manufacturers, such as Solectron. These companies often have a role that exceeds that of providing production capacity according to customer specifications and can encompass raw material purchases as well as product design and planning. Apple, for example, has an own facility for final assembly in Ireland. However, the company also relies on external vendors for final assembly, e.g. final assembly of all of Apple’s portable products (PowerBooks, iBooks, iPods) is performed by third-party vendors in Japan, Taiwan and China.

Supply chains in this industry consist of multiple fabrication, assembly, testing and packaging sites that are often located in different countries. The simplified supply chain depicted in Figure 1, for example, has two inbound chains, even though in reality there would be multiple. The upper one represents circuit boards, which are electronic components that are needed for final assembly. The chipset suppliers need to use sub-suppliers to obtain, for example, wafers. These wafer suppliers need other raw materials, such as silicon. These relationships are not represented in the figure. The lower inbound chain to final assembly (FAT) deals with other components, for example flips for mobile phones. The Other FAT supplier making those flips will need its own suppliers that deliver, for example, plastics (the 2nd tier supplier). Those plastics suppliers again need raw materials, such as plastic granules, which are made from oil or through reverse logistics processes from recycled plastics parts. Those upstream suppliers are not in the focus of the discussion in this paper, which is why they are not represented explicitly in the graph. In this work, the focus is on policy structures at the first tier suppliers, final assembly and testing and software flashing and packing stages.

Figure 1: Supply Chain with Two Sample Inbound Chains

Board Printing. The board printing (BP) stage is a fully automated process that is highly capital- and technology intensive. The process requires standard components, such as silicon wafers, as well as more complex components, such as integrated circuits and chipsets. These components are typically imported from other countries, which cre-

---

16 This simplified approach to visualizing a supply chain in the high-tech industry is also taken by Olhager et al.: Supply chain impacts at Ericsson, p. 47. See also Catalan and Kotzab: Assessing the responsiveness in the Danish Mobile phone supply chain, p. 672, for a more complex representation of the relationships. The definition of first tier suppliers, second tier suppliers etc. depends on the position of the supply chain node considered as the focal company. In this work, while the OEMs are focal companies, software flashing and packing and final assembly and testing are considered as the central part of the manufacturing process, even though both of these may be outsourced to other companies. The suppliers to final assembly are called the first tier suppliers in line with other representations of similar supply chains.
ates uncertain supply lead times.\textsuperscript{19} The board printing production step is typically a bottleneck process, since excess capacity is costly. While the equipment is capable of making a high variety of board designs, it is an economic necessity to operate the equipment at board printing at a high utilization rate.\textsuperscript{20}

\textit{Final Assembly and Testing.} The second stage of manufacturing is the final assembly and testing (FAT) process. This process is usually very manual, which is why it typically happens in China or other low cost countries, leading to relatively long delivery lead times. For example, in mobile phone production, this production step involves assembling the display and keypad into the product, as well as testing the device. Final assembly plants typically keep small stocks of component inventories whose size depends on the type of product and the material used.

\textit{Software Flashing and Packing.} Before the final software flashing and packing production step (SWF) there is a sizable inventory of semi-finished products, which is critical for short-term flexibility. The size of this stock is specified by the OEM and depends on the type of product. For example, stocks of standard mobile phones are typically larger than those of specialized mobile phones, such as phones co-branded with an operator, which are expensive to stock. At software flashing, after receipt of a customer order the software is flashed on the device, the manual, charger and other desired accessories are put into the box, and the package is then shipped to the customers, which are companies such as network operators or distributors.\textsuperscript{21} The reason that this step can only begin after receipt of a customer order is that the customers, for example mobile phone operators, often require customized products, with unique software and features that they expect to be installed by the OEM.\textsuperscript{22} This production step can happen in the same plant as final assembly and testing and is typically controlled by the OEM; for example, Nokia customizes the products at the final factors of distribution centre with a serial number (IMEI), software, and special features, such as logoed faceplates or special keypad buttons.\textsuperscript{23}

\subsection*{2.2. Difference between Push- and Pull-based Planning Approaches}

Differences in the characteristics of products and their markets create different requirements for supply chain management. In volatile markets, for example, supply chain capabilities may be required that are less important for products where demand is highly predictable. Also, these requirements may change over the product life cycle. Supply chain design therefore needs to consider the requirements of the products to be

\textsuperscript{19} Cf. Chang, Lin and Sheu: Aligning manufacturing flexibility with environmental uncertainty, p. 4769.
\textsuperscript{21} Cf. Olhager \textit{et al.}: Supply chain impacts at Ericsson, p. 47 and 49.
supplied via that chain. Different planning approaches may provide a means to tailor supply chains to different products.

The objective of supply chain planning is to match supply chain capabilities, such as available material and production capacity availability, with the demand characteristics faced by the supply chain. Planning is therefore concerned with developing plans and schedules that allow to capture sales opportunities and to satisfy customer needs in terms of speed, location and product variability. The decision rules in different planning approaches that relate to the activities necessary in supply chain planning, such as demand forecasting, inventory planning or materials management, have significant differences. In this context, pull systems, push systems and hybrid systems can be distinguished.

In a push-based supply chain, material processing is started when (1) material is available and (2) processing capacity of the next step is available. In other words, a push-based system is driven by a forecast as production and distribution decisions are based on long-term estimates of demand. In a push-based supply chain, assumptions are made about the system and its losses in order to attempt to match the final output with the customer requirements. Typical assumptions in push-based MRP systems include fixed demand and fixed lead times – if these conditions are not fulfilled, the original plan needs to be overridden. This can cause accumulations of inventory, for example at the suppliers with capacity limits.

In a pull-based supply chain, material processing is started when (1) material is available, (2) processing capacity of the next step is available and (3) an external trigger arrives. The external trigger would be the order from the immediate customer of the process, which can be both an internal or an external customer. In other words, the production and distribution processes in a pull-based system are driven by actual downstream demand and not forecasted demand. According to Ohno, who, among other things, devised a pull-based system for Toyota:

"Manufacturers and workplaces can no longer base production on desktop planning alone and then distribute, or push, them onto the market. It has become a matter of course for customers, or users, each with a different value system, to stand in the front line of the marketplace and, so to speak, pull the goods they need, in the amount and at the time they need them."
A pull-based system is simpler to implement than a push-based system and requires fewer interventions as the system reacts very quickly to small variations in demand. However, larger variations do require intervention by the planners as backlog situations are likely to occur. Another advantage of pull-based systems is the prevention of over-production and excessive inventory levels, which are core objectives of the lean philosophy, through an explicit limit of the amount of orders that can be in the system.\textsuperscript{28} Other advantages are that pull-based systems are simple and easy to understand and that all processes in the supply chain are synchronised with the customer.\textsuperscript{29}

A hybrid supply chain, also known as a push-pull system, combines the characteristics of pure push-based supply chains and pure pull-based supply chains. The initial stages of the supply chain, such as component suppliers, operate based on long-term forecasts, while the rest of the supply chain, for example the assembly and shipment processes, is driven by realized demand.\textsuperscript{30} In the high-tech electronics industry, most supply chains are hybrid supply chains, as the last steps of software flashing and testing are mostly performed after receipt of an actual customer order. However, the planning approach used for the rest of the supply chain may vary and could be either push-based or pull-based. This is represented in the two different planning approaches that are analysed in this work.

The different planning approaches may be more or less suitable for different types of products and at different phases of their life cycles. For example, Simchi-Levi, Kaminsky and Simchi-Levi argue that push systems are more suitable if demand uncertainty is low, while pull systems are more suitable if demand uncertainty is high.\textsuperscript{31} However, it remains to be analysed whether pull-based or push-based supply chains are suitable even for products with extremely high demand volatility. The impact of long component lead times on the appropriateness of the planning approach is a further issue of concern.

Some components needed for manufacturing high-tech products have long and variable manufacturing lead times. In the case of Cisco, components for networking products have lead times ranging between two and 24 weeks.\textsuperscript{32} Expensive application-specific integrated circuits (ASICs), for example, are unique to the respective product.\textsuperscript{33} In mobile phone production, for instance, each mobile phone model uses boards with different integrated circuits. Producing the wafers required to make the integrated circuits is a complex process characterised by long lead times, yield variations and capacity limitations.\textsuperscript{34} Equipment can also fail as machines break down. Also, the lead times in the supply chain and the output of the processes may vary and prices change.\textsuperscript{35} These

\textsuperscript{28} Cf. Hopp and Spearman: To Pull or Not to Pull: What Is the Question?, p. 142f.
\textsuperscript{33} Cf. Kok et al.: Philips Electronics Synchronizes Its Supply Chain to End the Bullwhip Effect, p. 39.
\textsuperscript{34} Cf. Gonçalves: Demand Bubbles and Phantom Orders in Supply Chains, p. 72–81.
\textsuperscript{35} Cf. Kaipia: The impact of improved supply chain planning on upstream operations, p. 1. For an example of how the interplay of limited capacity, short product life cycles, technological innovations and complex technology can create major price changes, see Helo, Petri: Managing agility and
long supplier lead times in the high-tech electronics industry are an important consideration for the selection of appropriate supply chain planning processes. Therefore, Christopher, Peck and Towill suggest a taxonomy based on lead times and demand predictability. For products with long replenishment lead times, they suggest lean supply chain design if demand is predictable, which would mean to “make or source ahead of demand in the most efficient way”. While this is necessary to some extent in the high-tech electronics industry due to the long component lead times, such an approach entails problems for supply chain management that need to be addressed with appropriate planning mechanisms. If demand is unpredictable – as typical is in the high-tech electronics industry – Christopher et al. suggest agile production and logistics postponement, where the final assembly steps are performed only after customer demand has been encountered. To some extent, the hybrid supply chains in the high-tech electronics industry already follow such a hybrid, “agile” strategy, as the software flashing and packing steps are only performed as actual demand is observed. The product specificity of some of the components, however, makes supply chain planning a challenge even if postponement strategies are implemented. This requires different approaches to design the supply chain efficiently and avoid costly inventory build-ups or stockouts.

The challenge for planners in the high-tech industry today, where companies sell a wide portfolio of products in volatile markets, is to design the planning approaches in the supply chain appropriately and implement them successfully. This is one possibility for achieving both increased efficiency and responsiveness in such supply chains as good planning approaches reduce the need for management attention. The scarce resources of management can then be dedicated to extraordinary situations, such as supply chain disruptions.

2.3. Overview of the System Dynamics Model

The System Dynamics model represents different planning systems that are conceived to be potential yet realistic approaches for supply chains in the high-tech electronics industry. Even though it may be theoretically possible to devise a system that incorporates all of the suggestions made by researchers on successful supply chain strategies, many of these are difficult to implement in practice. Typically, supply chains for high-tech electronics products, such as mobile phones or routers, are set up as hybrid push–pull production systems, with a varying extent of the use of each planning approach in the supply chain. The key distinction between different planning approaches in this supply chain model is achieved through differences in the decision-making policies by the supply chain members. Those locally rational decisions affect material or-

---


38 Cf. Kaipia and Holmström: Selecting the right planning approach for a product, p. 11.

ders, production levels and shipment rates in each period, and are based both on information known to the different players in the supply chain, e.g. a forecast provided to them, as well as on own beliefs about future developments. Due to aspects of bounded rationality, such policies can lead to unexpected dynamic behaviour in the system.\textsuperscript{40}

The first supply chain planning approach that is modelled is a push-based supply chain planning approach, in which a central planning department prepares the forecasts for components. This push approach, based on long-term demand forecasts, is used by the supply chain members upstream of software flashing, up to a major buffer stock of semi-finished products. This buffer stock is located at the software flashing and packing production step. From there, the customer pulls the product and the final customization steps and shipment are performed within a very short lead time and driven by customer orders. In the mobile phone industry, for example, operators, such as T-Mobile or Vodafone, place their orders directly with the country representatives of the mobile phone OEMs, such as Nokia. Cisco calls such an approach “build to stock and ship to order”.\textsuperscript{41}

The second planning approach is a pull-based system from the customer until the board printing stage. Suppliers to board printing, e.g. semiconductor chip manufacturers, still use a push approach, since the lead time for these components is very long, i.e. several weeks up to months. Each of these modelled planning approaches can receive several different demand patterns as an input; in this paper, a typical demand pattern for a high-tech electronics product is analysed. The flexibility of the model allows identification of the most appropriate approaches for the supply chain planning and ordering mechanisms for each of several demand scenarios with different demand uncertainty.

The underlying supply chain set-up is based on the structure of a typical high-tech supply chain and currently represents decision policies for four stages plus the customer. The structure represented in the System Dynamics model consists of a customer, software flashing and packing, final assembly and testing, and first tier suppliers. In addition to board printing, a further supplier to the final assembly stage is also explicitly included in the model. The model considers the information and material flows through the supply chain, beginning with the initial customer order for a single product and ending with its delivery. This is done with the objective of obtaining a fundamental understanding of the underlying structure of the decision-making processes at the different supply chain nodes. Considering fewer supply chain tiers would have been a serious limitation of the research. Figure 2 provides an overview of the model structure. At each step in the supply chain materials are converted to finished goods (FG), while the finished goods at the supplier and board printing level then become inputs for the production process at the final assembly stage, and the finished goods at final assembly become inputs for the customization process at software flashing. At each of these echelons, there are thus materials, work in process and finished goods inventory levels.

The supply chain is adaptable to different conditions through various parameters. These include delays in information processing, production and shipment delays and resulting lead times, capacity limitations and external shocks to the supply chain that can happen at board printing and at the supplier. Most of these parameters do not differ


\textsuperscript{41} Parmar: Supply Chain Architecture in a High Demand Variability Environment, p. 12.
for comparing different planning approaches, which enables a comparison of performance measures purely based on changes in decision-making policies. From this generalized structure, insights can be gained into the dynamic performance of the system and opportunities for improvement.

![Figure 2: Overview of the Structure of the System Dynamics Model](image)

### 2.4. Decision Rules in the Two Planning Approaches Represented in the Model

#### 2.4.1. Push-Based Planning Approach with Central Planning

In an approach that is primarily based on push principles with central planning up to the software flashing production step, there is one global plan. This plan drives production at the upstream suppliers. Long supplier lead times for chipsets needed for board printing (e.g., 46 days) require that second tier suppliers base their production and shipments on a forecast. Such a forecast is initially prepared by the Original Equipment Manufacturers (OEMs), which typically run the software flashing part of the supply chain. This initial forecast is based on information from the sales people, who talk directly to customers and then aggregate the forecasts to product groups. The planners at software flashing adjust this forecast to accommodate deviations of demand expectations and actual demand and to keep their raw materials inventory at the desired level. Then they communicate this adjusted forecast to the final assembly and testing stage. Once the product is released to the market, this forecast prepared by software flashing is also adjusted based on the difference between the original forecast and actual customer demand to account for forecast errors. The planners at the software flashing stage can additionally place emergency orders with the chipset suppliers. The chipset suppliers may be able to supply a limited amount of chipsets faster than usually – and typically at a higher cost. Anything that is produced by the suppliers based on the global plan is then pushed through the other production steps, board printing and final assembly, until the products reach the buffer stock at the software flashing stage. This approach is visualized in Figure 3.
Software Flashing and Packing (SWF). The processes at the software flashing and packing stage are driven entirely by customer demand. Software flashing begins performing the final software flashing and packing steps nine days before the customer request date (CRD), which corresponds to the time it takes to perform these steps and ship the product to the customer. To trigger this process, customer demand for any given day is received by software flashing nine days ahead of that day.

Figure 4 visualizes the order fulfilment process at software flashing. Orders are received nine days ahead of time as incoming production requests SWF and flow into a stock of Backlog SWF, which represents the amount of orders waiting to be produced and is depleted by production starts. The number of products whose production begins in a given period is determined by (1) the desired orders in production SWF, which are the desired amount of orders that should begin processing at software flashing, and (2) the availability of sufficient raw materials inventory (Inventory RM SWF). The value of desired orders in production is calculated by taking the new incoming production requests, i.e. the orders that are supposed to begin processing at software flashing because of incoming orders in this period, plus any orders that may have accumulated in the Backlog SWF. The equation is:

\[
\text{desired orders in production SWF} = \text{MAX}(0, \frac{\text{Backlog SWF}}{\text{minimum time to start production SWF}} + \text{incoming production requests SWF})
\]

The Feasible production SWF is then defined as the amount of desired orders in production SWF or possible production given the amount of raw materials available, whichever is smaller:

\[
\text{feasible production SWF} = \text{MIN} \left( \text{desired orders in production SWF}, \frac{\text{Inventory RM SWF}}{\text{materials used per FG unit SWF}} \right)
\]

The variable materials used per FG unit SWF is included to allow possible future modifications of the model with more than one unit of raw materials used per finished good unit, and is set to 1. The minimum time to start production SWF is also set to 1.
Feasible production SWF =
\[ \text{MIN}\]
\[ \left( \text{desired orders in production SWF,} \right. \]
\[ \text{Inventory RM SWF/materials used per FG unit SWF} \]
\[ \left. /\text{minimum time to start production SWF}\right) \]

Once production at software flashing is started, the stock of Backlog SWF is reduced by the amount just calculated.

![Diagram]

Figure 4: Push: Order Fulfilment SWF

The production process at software flashing, depicted in Figure 5, is represented by a third-order exponential delay of the Feasible production SWF that converts materials into finished goods, flowing into the stock of Inventory FG SWF, which is the stock of finished good as software flashing that is immediately shipped to the customers. The pipeline stock of work in process inventory is also computed to allow inventory cost calculations. As soon as the production process is finished and products reach the finished goods inventory they are shipped to the customer. The customer then receives them after a third-order exponential shipment delay of four days.
In the planning approach with central planning, a single central plan is prepared and communicated to the first tier suppliers. This plan is prepared by the central planners that are based at the software flashing and planning stage, which corresponds to the OEMs. Before the product is introduced to the market these forecasts are intended to determine initial staging of components at second tier suppliers, i.e. (1) the supplier to the board printing stage and (2) the supplier to the first tier component supplier. Due to the supplier lead times, the forecast for board printing needs to have a horizon of 76 days, while that for the other supplier has a forecast horizon of 36 days. The original base forecasts for customer demand are read as input data from an Excel spreadsheet. This is necessary because before the product is introduced to the market the forecast needs to be based on information that is exogenous to the model. After the product introduction this forecast is adjusted by software flashing based on two factors, (1) an adjustment to account for the difference between the base forecast and actual demand once the product has been released to the market and (2) an adjustment to the forecast aimed at achieving a specified goal for the desired level of raw materials inventory at software flashing. These adjustments can be switched on and off such that their impact on the system can be analysed. The overall structure of the forecast adjustment process at software flashing is represented in Figure 6.
Finally, software flashing can place emergency orders with alternative chipset suppliers that have shorter delivery delays. These emergency orders are based on the size of the backlog observed by software flashing. Similarly to the adjustment for inventory, the amount of emergency orders is calculated as a goal-adjustment process, with the goal being a backlog at software flashing of zero units.

**Final Assembly and Testing (FAT).** In the push-based planning approach with central planning the production process at the final assembly and testing stage is driven by the amount of production that is feasible based on the raw materials that are available, i.e. there is no forecast that drives production and production continues as long as enough raw materials, i.e. both boards and other supplies, are available. This is visualised in Figure 7 and corresponds to pushing the parts through the supply chain, following the central plan that drives production at the suppliers.
First Tier Suppliers (BP/Sup). In the push-based planning approach with central planning, the first tier suppliers, i.e. board printing and the other supplier to final assembly, receive the central plan and order the materials from their raw materials suppliers based on exactly this plan. Board printing orders the forecast through the standard channel and considers emergency orders separately, while the supplier orders the forecast as well as the emergency orders from the same second tier supplier. Production at the first tier suppliers is then, similarly to that at final assembly, driven by the amount of materials they receive from their suppliers.

2.4.2. Pull-Based Planning Approach

In a pull-based planning approach, material orders are based on actual demand and on changes to safety stock levels. Specifically, in each period a supply chain member orders materials to fulfil demand of a period, as well as a potential additional adjustment to achieve the desired safety stock level. This approach is visualized in Figure 8. For the different supply chain tiers modelled, i.e. software flashing and packing, final assembly and testing, and the first tier suppliers, some details of these decision rules are explained in the following.

---

Figure 8: Pull Loops

Software Flashing and Packing (SWF). The production and shipment processes at software flashing as well as the preparation of the central plan are identical to those in the push-based planning approach with central planning. A central plan is necessary even in a pull-based planning approach since the first tier suppliers need to place orders with their suppliers according to a long-term forecast. The difference of the pull-based approach lies in the policies regulating the receipt of materials from final assembly. In contrast to the push approach, the pull approach triggers a shipment from final assembly to software flashing only when an order by software flashing has been placed. This ordering logic is based on a pull replenishment logic that is explained in the following.

First, the number of orders to be placed with a specified order size is calculated. This number is then multiplied with the order size resulting in the amount of materials to be ordered each day. This value is accumulated in the stock Materials on order SWF that itself is reduced by the material arrival rate at software flashing. There is also a possibility to limit the amount of orders placed per day to a maximum value. In the base case, however, this limit is not used.

The number of orders to be placed is calculated as follows. The material orders itself consist of several factors that are summed up. First, the incoming production requests SWF, which represents customer demand in the current period, are part of material orders. Considering new incoming orders is, however, not sufficient for driving the ordering process. Therefore the forecasted demand during the lead time until any orders are received from FAT also becomes part of the calculation, since this period needs to be covered with materials on order. This forecast for demand during the expected lead time fluctuates if demand is different from the original forecast, as the forecasts used for the calculation are the adjusted forecasts. Then, the safety stock size is added to the material orders. Also, the backlog at SWF is considered since orders need to be placed to consider any orders that could not yet be fulfilled. From this, the inventory position, i.e. the amount of materials on order plus the current inventory level at SWF, is deducted. The inventory position represents the amount of materials already ordered or received, and
creates a limit to the amount of orders that are in the system.\textsuperscript{44} The INTEGER function is finally used to identify the number of orders to be placed in each period when dividing the desired orders through the order size.

\begin{verbatim}
no of orders SWF =
   INTEGER
   {
   (incoming production requests SWF
   * materials used per FG unit SWF
   * time to place an order SWF)
   + forecast demand in materials LT SWF
   + safety stock size SWF
   + initial stock SWF
   + (Backlog SWF * materials used per FG unit SWF)
   - Materials on order SWF
   - Inventory RM SWF
   )
   /order size SWF
   /time to place an order SWF
   }
\end{verbatim}

The safety stock size at software flashing is calculated as the standard deviation of demand over the material lead time, which is 10 days, multiplied with a safety factor $z$, representing the number of standard deviations of demand to be covered. In the base scenario this safety factor is set to 1.65, which corresponds to a service level of 95 percent. The standard deviation of demand is calculated every 10 days using an algorithm developed by John D. Sterman.\textsuperscript{45}

In addition to the safety stock size $SWF$ variable there is also the possibility to include an additional initial stock $SWF$. This initial stock can incorporate planned product introductions in the model even when currently no demand is observed, i.e. before the product introduction. Except where otherwise indicated, to simulate a pure pull-based system based solely on realized demand this initial stock is zero.

\textsuperscript{44} See Hopp and Spearman: To Pull or Not to Pull: What Is the Question?, pp. 142–143.
Final Assembly and Testing (FAT). Production and shipment processes at final assembly in the pull planning approach begin only after orders from software flashing have been received. Material orders are placed following the same policies as at software flashing. The incoming orders, i.e. current period demand, are the incoming order rate FAT, which is the sum of normal material orders placed by software flashing at final assembly and emergency orders.

First Tier Suppliers (BP/Sup). Production and shipment processes at board printing in the pull planning approach begin only after orders from final assembly have been received. Standard material orders, however, are placed according to a forecast prepared by central planning. The orders are adjusted to account for the difference between that forecast and actually experienced demand, i.e. orders placed by final assembly, and for a desired inventory level at board printing. This desired stock level is determined by a desired number of days of stock based on average demand observed from final assembly. Emergency orders by software flashing are placed separately as in the other two planning approaches.

Similarly to board printing, production and shipment processes at the other first tier supplier in the pull planning approach begin only after orders from final assembly have been received. Material orders, including emergency orders, are placed according to a forecast prepared by central planning. This forecast is adjusted to account for the difference between that forecast and actually experienced demand, i.e. orders placed by final assembly, and for a desired inventory level at the first tier supplier. The desired inventory level is is determined by a desired number of days of stock as determined by average demand from final assembly.
3. Evaluation of Planning Policies to Achieve Responsiveness and Efficiency for a Product with a Typical Product Life Cycle

Product life cycles in the high-tech industry are characterised by a strong increase in demand during the introduction phase, followed by a gradual downward trend during maturity and finally an end-of-life demand drop, often caused by a new product generation being introduced. The demand pattern for the simulations is the product life cycle shown in Figure 10, with an initial sales peak of close to 17,000 units on day 130 followed by a gradual downward trend in demand and a steep drop at the end of the product life cycle. The initial forecasts predict exactly this product life cycle.

![Figure 10: Demand Pattern]

In reality, demand is subject to variation over time. Such demand volatility is introduced in the model in order to analyse its impact on the performance of the different planning approaches. Three demand scenarios are analysed – one scenario without random demand variations, one scenario with a low amount of variation, and one scenario with a high amount of variation. The forecast in each of these scenarios is the product life cycle shown above, without noise, and mean demand is also corresponding to this product life cycle. Noise is added to the previously stable product life cycle through a pink noise process, which is first-order autocorrelated noise where the next random variation depends in part on the previous variations. This is necessary to represent that one period’s demand is not independent of the last but depends to some degree on history, i.e. there is some inertia. In the simulations with random noise, the supply chain performance measures discussed refer to the mean values obtained from a sensitivity analysis with 200 different random demand patterns. The values indicated for inventory turns and delivery performance to customer request are the mean values over 200 simulations with different noise seeds, i.e. 200 demand patterns with different noise patterns, each with mean demand following the product life cycle shown above. The standard

deviation of the pink noise remains constant over these 200 simulations and is 0.1 in the low noise scenario and 0.8 in the high noise scenario.

Driven by the variation in demand the planners in all planning approaches are confronted with the same problem – how to satisfy customer demand on time considering that the raw material receipts are unlikely to constantly coincide with the requirements during a particular period. If demand is below the forecast for a few days, inventory will build up, and if demand is higher than the forecast, the inventory levels will decline until eventually a backlog situation arises as demand can no longer be fulfilled.

3.1. Effect of Random Demand Deviations

The push approach achieves the highest level of delivery performance in each of the three demand scenarios with different demand volatility. This approach achieves 86 percent in the demand scenario without noise, 74 percent in the low noise scenario, and 41 percent in the high noise scenario. While the pull-based planning approach outperforms the push-based planning approach on efficiency in the scenarios with low noise and no noise, its delivery performance is consistently lower than that achieved by the alternative planning approach. Delivery performance in the pull-based planning approach is 60 percent in the no noise scenario, 56 percent in the low noise scenario, and 24 percent in the high noise scenario. Depending on the objectives of the organisation, the lack of planning for the initial demand spike and consequential high initial backlog makes the pure pull-based planning approach unattractive.

Figure 11: Sensitivity to Randomness in the Demand Pattern
Table 2: Sensitivity to Randomness in the Demand Pattern

<table>
<thead>
<tr>
<th>Product Life Cycle</th>
<th>Base Case</th>
<th>Low Noise&lt;sup&gt;48&lt;/sup&gt;</th>
<th>High Noise&lt;sup&gt;49&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Delivery Perf.</td>
<td>Inventory Turns</td>
<td>Delivery Perf.</td>
</tr>
<tr>
<td>Push</td>
<td>86%</td>
<td>14.0</td>
<td>74%</td>
</tr>
<tr>
<td>Pull</td>
<td>60%</td>
<td>16.2</td>
<td>56%</td>
</tr>
</tbody>
</table>

3.2. Planning the Product Introduction in the Pull-Based Planning Approach

A major problem of the pull-based approach is the low delivery performance in the initial phase of the product life cycle. In order to accommodate this difficulty, the product introduction is planned for by initialising production before customer demand is observed. Software flashing and final assembly and testing therefore order materials with the objective to have an initial stock of 100,000 units of components available when the product is introduced to the market.<sup>50</sup> This intention is also known to the suppliers at the beginning of the forecasting process. In period 120, i.e. 20 days after the product introduction, this desired amount of materials in stock to cover for the product introduction phase is reset to zero.

For the demand scenario with high demand uncertainty the initial staging of the supply chain through pre-production of components and stocking these at software flashing and final assembly increases average delivery performance to 40 percent (from 24 percent) and increases inventory turns to 9.0 (from 8.4). Delivery performance is equal to that in the push-based planning approach and efficiency as measured by inventory turns is slightly lower.

---

<sup>48</sup> Mean values over a 200 simulations with different noise seeds; standard deviation of white noise used for pink noise process is 0.1; this noise is added to the constant demand. Pink noise is first-order autocorrelated noise where the next random shock depends in part on the previous shocks.

<sup>49</sup> Mean values over a 200 simulations with different noise seeds; standard deviation of white noise used for pink noise process is 0.8.

<sup>50</sup> In the model, each final product at software flashing is composed of one finished good unit received from final assembly. Similarly, each finished good at final assembly is composed of one finished good unit from each first tier supplier and each finished component at the first tier suppliers is composed of one unit of components received from the respective second tier supplier.
The stability of the forecasts provided to the component suppliers in this approach is also high, which is beneficial for the supply chain. This can be seen when comparing the forecasts sent to board printing in the pull-based and push-based planning approach in Figure 13 to the initial forecast corresponding to the expected product life cycle.

Figure 13: Forecast for Board Printing (example, not averaged)

This figure shows the performance of incorporating the base safety stock level into the initial forecast, which leads to a delivery performance of 75 percent and inventory turns of also 7.2 for the pull-based planning approach with a base safety stock level.
A major drawback of the pull-based planning approach is the large degree of variation in the production rate at the suppliers, which is caused by the large variability in material orders by final assembly. As can be seen in Figure 14 for a sample simulation run, production of components in the push-based planning approach with central planning follows the actual product life cycle relatively closely, even though production may be ramped up to the full capacity at board printing. This eliminates a large degree of the variation of end customer demand, which is beneficial to the efficiency of the supply chain as high and relatively stable capacity utilization is a success factor for efficient processes at board printing. In contrast, the production rate at board printing in the pull-based planning approach experiences large fluctuations with subsequent periods of both production at the capacity limit as well as production at zero. This is caused by the direct transmission of demand data through the supply chains, where relatively large deviations in demand cause relatively large deviations in material orders, even if these are only issued in large batches. Such large variations can pose a major challenge to the production planners at the suppliers, as it may be difficult to ramp up and ramp down capacity within these relatively short time intervals.

![Figure 14: Production at Board Printing (example, not averaged)](image)

### 4. Achieving Responsiveness and Efficiency over the Product Life Cycle

The decisions taken to achieve an appropriate level of both dimensions of supply chain performance, efficiency and responsiveness, need to take into consideration aspects related to the product life cycle. Sales in the initial phase of the product life cycle are particularly important, as higher prices can be charged and product life cycles are often very short. Therefore, high delivery performance is particularly desirable during the initial phase of the product life cycle. For the typical product life cycle of a high-tech product as modelled in the simulations the time behaviour of delivery performance over time varies significantly between the different planning approaches. As an example Figure 15 shows the demand scenario with high noise and the resulting delivery per-
formance for the different planning approaches. The push-based planning approach with central planning achieves relatively high delivery performance at the beginning of the product life cycle, ranging in the example between 100 percent and around 50 percent. The pull-based planning approach in its pure form, however, is unable to deliver on time from the beginning of the product life cycle – the first unit is delivered on time around day 250 – and therefore a large amount of sales are potentially lost as customers move to competitors.

![Figure 15: Delivery Performance (example, not averaged)](image)

It has been shown that planning for the product introduction by producing a certain amount of products before customer demand is first observed can increase delivery performance with only a small reduction in efficiency. As illustrated in Figure 15 for the product life cycle demand scenario with high demand volatility, such an initial stock level allows a delivery performance that exceeds that of the push planning approach with central planning. The range of delivery performance in the example is between 75 percent and 100 percent in the first weeks after the product introduction. At the same time a value of inventory turns of 9.0 on average is achieved, which is only slightly worse than in the push planning approach, achieving 9.6 on average.\(^{52}\) In addition, both the push-based and the pull-based planning approach with initial staging of the supply chain are also able to react to supply chain disruptions in a meaningful way, such that delivery performance only experiences a small reduction. On average, delivery performance is only reduced by one percentage point if production at board printing is disrupted for 10 days in both planning approaches.

A major concern with regard short product life cycles is inventory of products and components that is unsold at the end of the product life cycle. Due to the characteristics of the two push-based planning approaches, both are subject to major inventory build-ups at all stages of the supply chain both during and at the end of the product life cycle.

\(^{52}\) As in the preceding discussion, an initial stock of 100,000 units is introduced for the raw materials inventory levels at software flashing and final assembly. In addition, the initial forecast is adjusted to consider these planned stock levels for the product introduction phase by a one-time peak of 200,000 units plus the original forecast in the first forecast period.
Inventory build-ups are particularly detrimental downstream in the supply chain due to the higher value of the products as they near completion. Such downstream inventory build-ups, as well as inventory build-ups at final assembly, are avoided in the pull-based planning approach. Due to the different decision making policies in the pull-based planning approach, with no adjustments for desired inventory levels, inventory of semi-finished products at software flashing at the end of the product life cycles is significantly lower than in the other planning approaches, independent of the demand pattern. However, inventory build-ups in the pull-based planning approach may still occur at the upstream suppliers, which are driven by a forecast.

There, however, the cost of inventory is significantly lower, which also reduces the damage if obsolete inventory needs to be discarded. In the demand scenario with the second product life cycle, perfect forecasts and high noise, for example, the raw materials inventory levels at software flashing at the end of the simulation in the push-based planning approach with central planning correspond to 420,000 Euros, and in the pure pull planning approach they are only 45,000 Euros. In the pull approach with initial staging of components for the first phase of the product life cycle, the end raw materials inventory level at software flashing in this demand scenario is worth 50,000 Euros on average. At the first tier suppliers, mean inventory levels at the end of the simulations are 5,000 Euros (Push), 270,000 Euros (Pull) and 180,000 Euros (Pull with initial stock). To put these values into perspective, consider that total sales over the product life cycle are 3,000,000 units, with each unit valued at 1 Euro. Inventory levels at the end of the product life cycle can therefore, on average, reach up to almost 20 percent of total sales. The pull approach that plans for the product introduction can therefore achieve a higher level of delivery performance in the initial phase of the product life cycle with lower average inventory levels in the supply chain at the end of the simulation. Overall inventory turnover over the course of the product life cycle, however, is slightly lower than in the push-based planning approach. The results in all of the other simulations with different levels of demand volatility are comparable.

![Average Inventory Levels at the End of the PLC](image)

**Figure 16: Average Inventory Levels at the End of the Product Life Cycle**

---

53 All values are averages over 200 simulations with different noise seeds and rounded to full 5,000.
It has been shown that it is possible to design the planning policies in the high-tech electronics industry in such a way that the performance of the supply chain is improved on both dimensions, responsiveness and efficiency. A combination of a simple pull-based planning system with a plan for the initial phase of the product introduction can perform as well as or even better than a push-based planning approach with central planning, independent of the demand characteristics of the product. This finding has a number of implications. First, it shows that both objectives, responsiveness and efficiency, can be achieved in a supply chain with one planning approach. Segmenting the product portfolio may not be necessary.

The finding that pull-based planning approaches can be used even in environments with high demand uncertainty and long component lead times is valuable if one considers issues of implementation. Push-based planning approaches are more difficult to implement than pull-based planning approaches, in particular if the supply chain crosses multiple organizational boundaries. Planners in push-based planning systems depend on information about capacities, inventory levels, and generally the structure of the supply chain. For example, if information about supply chain disruptions at the supplier side is not communicated to the supply chain planners, as in this model, then their observations of such disruptions will only occur after a significant time delay.

In contrast, a pull-based system needs much less information and depends entirely on local decision rules at each supply chain echelon, which would immediately adjust to supply chain disruptions, for example, through local decision heuristics. The only act of planning that is needed is the long-term forecast that is communicated to the suppliers. As has been shown, these forecasts can be more stable in a pull-based planning system, which is advantageous for the supply chain.

However, there are also notable drawbacks to implementing the pull-based planning approach in the high-tech electronics industry. A complication when implementing pull-based planning systems in such an environment with a high potential for demand variability is that the production rate at suppliers may be subject to large fluctuations if these are integrated in the pull-based planning system. These fluctuations may require higher average inventory levels at the suppliers. If there were several product variants, however, which use some of the same base components and could be produced on the same line at board printing, then this would smooth the overall production rate. However, producing different variants on the same line would also add additional complexity as the production of the different variants needs to be scheduled to coincide with the demand. Correlations between different variants also need to be considered. The feasibility of such a system therefore depends on the product characteristics in terms of mix variability as well as the flexibility of changing production volumes at board printing. Also, the suppliers may need to be financially compensated for larger variations in demand and resulting increases and fluctuations of their raw materials inventory levels.
References


