Comparison between production controls in multi-stage multi-product manufacturing environments: two case studies

J. Prakash\textsuperscript{a*} and J.F. Chin\textsuperscript{b}

\textsuperscript{a}Faculty of Engineering and Green Technology, Universiti Tunku Abdul Rahman (UTAR), Jalan Universiti, Bandar Barat, 31900 Kampar, Perak, Malaysia; \textsuperscript{b}School of Mechanical Engineering, Universiti Sains Malaysia (USM), Engineering Campus, 14300 Nibong Tebal, Penang, Malaysia

(Objective): The paper presents a study on parallel CONWIP systems based on two case studies benchmarked against CONWIP and push. The performance measures are flow time and work-in-process (WIP) level.

(Methodology): A discrete event simulation model was constructed and was used to run experiments for individual production controls. The gathered data were analyzed through response surface methodology.

(Results): In case study 1, the parallel CONWIP with FIFO dispatch rule is superior at both flow time and WIP level. In case study 2, push, CONWIP, and parallel CONWIP with LH dispatch rule to be considered superior at both flow time and WIP level.

(Conclusion): One notable advantage of parallel CONWIP systems is the ability of the card count to respond to changes in the demand for product classes. Several future research directions are proposed, such as framework of dynamic card controlling, and implementation of parallel CONWIP aligned with lean practices.

(Keywords): CONWIP; parallel CONWIP; production control; reduce flow time; reduce WIP

1. Introduction

Manufacturers today can no longer follow the path set by Henry Ford, that is, market shares and high profits are captured by producing large volumes of a standardized product. This shift toward a larger product variety emanates from factors such as the globalization of markets, the growing sophistication of customer preferences, and the changes in trading structure (MacDufﬁe, Sethuraman, & Fisher, 1996). Bils and Klenow (2001) studied the increase in product variety offered by various industries and concluded that, on average, product variety increases by 1% each year. Product variety growth rate is, in turn, dependent on various factors, such as the distance from the frontier of export variety and the intensity of research and development (Addison, 2003). This claim is supported by manufacturing facilities performing several alterations in their production process to support the increase, such as a reduction in product development time, minimization of resource consumption, and utilization of technologies capable of product customization (Alford, Sackett, & Nelder, 2000; Denton, Gupta, & Jawahir, 2000;...
These efforts have contributed to the reduction of cost in offering a broad variety of products, enabling firms to offer greater customer value and increase profit (Gao & Hitt, 2004).

Key decisions in product variety expansion can generally be classified into product variety creation and product variety implementation (Lancaster, 1990). Product variety creation decisions focus on the quantity, variety, and timing of product introduction to target markets, and product variety implementation decisions focus on the strategy employed by a manufacturing organization to implement the variety creation decisions (Ramdas, 2003). One crucial product variety implementation decision is the installation of a suitable production control. Production control is defined as a set of activities involved in handling materials during manufacturing, taking into account the demand, resource availability, cost, and capacity constraints (McKay & Wiers, 2004). The objective of production control is to achieve the required quality and quantity of products within the time needed. In the field of operations management, the term ‘multi-stage multi-product,’ which was coined by Villa (1989), is a formalization frequently used to address the various architectures of the production process consisting of more than one processing stage and more than a single product type. Production control can be classified into push or pull systems (Karmarkar, 1991). In a push system, an upstream workstation begins processing without a request from the workstation downstream. In a pull system, an upstream workstation begins processing when it is triggered by the workstation downstream.

A production control in a multi-stage multi-product environment encounters two challenges (Smalley, 2009): (a) fulfilling the demand for product variety and (b) maintaining low inventory. First, the demand for product variety in a multi-stage multi-product environment ranges from low to high. For products with high demand, demand forecasts are generally available because they can be estimated from historical data. In this category, work-in-processes (WIPs) are maintained for frequent consumption and replenishment to ensure that the demand is continually met. Demand forecasts may not be available for products with low demand. In this category, WIPs are usually not kept as less frequent consumption and replenishment may result in their deterioration over time. Second, push systems, although simple, may result in high WIP level and long flow time (Karmarkar, 1991). On the other hand, pull systems are important in maintaining low inventory, particularly WIPs, as they regulate their movement. Various methods to limit the WIP level can be used, and the chosen method must facilitate the degree of product variety offered by the manufacturing facility (Framinan, González, & Ruiz-Usano, 2003).

Production controls in multi-stage multi-product environments have seen a shift from the traditional push to pull systems, as evident from the number of studies comparing their performances, with the latter deemed superior (Khojasteh-Ghamari, 2009; Lavoie, Gharbi, & Kenné, 2010). Many pull systems are continuously developed or modified to cater more effectively to different environments. This paper proposes a pull system in multi-stage multi-product environments and compares it with relevant production controls using two case studies. The objective is to determine the best production control that yields the shortest flow time and lowest WIP level. This paper is organized as follows: Section 2 provides a review of the production controls in multi-stage multi-product environments and the common alterations of constant work-in-process (CONWIP) systems. Section 3 formally describes the generalized model of this study. Section 4 explains the adopted methodology, which involves discrete event simulation through two case studies. Sections 5 and 6 explain the first and second case studies, respectively,
including the background, models, results, best production control, and optimal operating parameters. Section 7 discusses the findings from both case studies, and Section 8 concludes the paper.

2. Literature review

Push systems in a multi-stage multi-product environment commonly employ a production planning system known as manufacturing resource planning (MRP II), which works backwards from a production schedule of finished goods to derive schedules for raw materials, taking into account the WIPs and finished goods inventory. MRP II overcomes the limitations of traditional material requirements planning (MRP) by tracking inventory while managing the schedule. Today, MRP II has been developed to become enterprise resource planning (ERP), which integrates the financial, material, asset, and resource aspects of a manufacturing organization (Schutt, 2004). Nevertheless, irrespective of MRP II or ERP, the planning system functions as a ‘push’ trigger that pushes products downstream based on the schedules of the raw materials (Hopp & Spearman, 2001). A push system is based on the premise that raw materials are released at scheduled times and pushed from one workstation to another based on that schedule (Lyons, Mondragon, Piller, & Poler, 2012). Many studies that discuss push systems in multi-stage multi-product environments can be found in the literature and range from analytical studies (Barba-Gutiérrez & Adenso-Díaz, 2009; Mohammaditabar, Ghodsypour, & O’Brien, 2012; Mula & Poler, 2010; Robinson, Sahin, & Gao, 2005) to implementation procedures (Kanet & Stößlein, 2010).

The four pull systems used in multi-stage multi-product environments are kanban, CONWIP, paired-cell overlapping loop of cards with authorization (POLCA), and mixed pull. In the kanban system, production is initiated through the progressive transfer of cards to upstream workstations, but the circulation of cards between two consecutive workstations controls the WIP level between them (Wang, 2010). Studies on kanban systems in multi-stage multi-product environments are mainly analytical (Gurgur & Altıok, 2008; Krieg & Kuhn, 2008). Kumar and Panneerselvan (2007) reviewed kanban systems and found that kanban requires product variety to be kept at a minimum because of the repetitive nature of production. A mixed pull system provides a more practical approach, in which a kanban system is used only for make-to-stock (MTS) products and a push system is used for remaining make-to-order (MTO) products (Smalley, 2009).

One pull system commonly used in multi-stage multi-product environments is the CONWIP. Kanban systems generally provide localized WIP control through card circulation between two consecutive buffers, whereas the CONWIP system limits the WIP level in a line through card circulation between the input buffer and finished goods buffer (Spearman, Woodruff, & Hopp, 1990). The literature addressing CONWIP systems in multi-stage multi-product environments ranges from analytical studies (El-Khouly, El-Kilany, & El-Sayed, 2009; Lee & Chen, 1997; Mönch, 2005; Parvin, Van Oyen, Pandelis, Williams, & Lee, 2012; Rose, 2001; Yang, Hsieh, & Cheng, 2011) to implementation frameworks (Li, 2009; Slomp, Bokhorst, & Germs, 2009).

In a POLCA system, a received card represents a capacity vacancy in a cell, which causes localized WIP control between two consecutive cells (Krishnamurthy & Suri, 2009). Research on the POLCA system has mainly been analytical studies (Germs & Riezebos, 2010; Riezebos, 2006) and implementation frameworks (Araya, 2012; Krishnamurthy & Suri, 2009; Vandaele, Nieuwenhuyse, Claerhout, & Cremmery, 2008).
In a mixed pull system, product variety is classified into three classes: high-runner (HR), medium-runner (MR), and low-runner (LR). HR, MR, and LR contain product families with high, medium, and low demand, respectively. Moreover, the product families in HR, MR, and LR have low, medium, and high variability in quantity, respectively. HR and MR use a common kanban system, and LR uses a push system. Case studies were mostly conducted in studies on mixed pull system in multi-stage multi-product environments (Bar, 2006; Gates, 2004; Horbal, Kagan, & Koch, 2008; Saurin, Marodin, & Ribeiro, 2011; Skelley, 2009).

Push systems have one distinct advantage over pull systems. A pull system can only be effective if the operational parameters are correctly selected and the functional conditions are fulfilled (Lee, 1989). For instance, a kanban and mixed pull system requires setting the card count for HR and MR, proper machine layout, short setup time, and standardization of work (Monden, 1983). A CONWIP system requires setting a single card count for the line and an allowance for work ahead (Spearman et al., 1990). A POLCA system requires reconfiguration of workstations into cells, setting card count in each cell loop, and knowledge of production flow time in each cell (Krishnamurthy & Suri, 2009). A push system only requires the availability of raw materials, which are released at scheduled times based on the schedule by the planning system (Lyons et al., 2012).

Pull systems are favored over push systems. Pull systems have less congestion and, thus, a smaller production flow time than push systems (Spearman & Zazanis, 1992). Second, controlling the WIP level (pull systems) is easier than controlling the throughput (push systems) because WIPs can be observed directly but throughput can only be estimated from the capacity and production flow time (Hopp & Spearman, 2004). Inaccurate estimations could cause target production to be exceeded or missed (Schutt, 2004). Third, the production flow time variances are smaller in pull systems than in push systems, resulting in the better estimation of production lead time and safety stock required (Hopp & Roof, 1998).

Kanban, mixed pull, and POLCA systems require the fulfillment of the various conditions mentioned. Less pristine environments may face difficulties in fulfilling all conditions. For instance, in attaining a proper machine layout and reconfiguring workstations into cells, machines in a facility may have to be relocated or reconfigured extensively, resulting in the stoppage of production during the transformation period (Harrison & Petty, 2002). A short setup time calls for a close study of the setup process and may require extensive use of costly technology to modify the machine (Mahadevan, 2010). Standardization of work consumes a substantial amount of time and is difficult to achieve because the preconditions, such as cyclical motion, limited equipment downtime, or lack of quality problems, cannot be met (Kató & Smalley, 2011). CONWIP systems are more pertinent as conditions are less rigid. Spearman and Zazanis (1992) proved that the success of a pull system owes more to the limitations of a WIP level within a boundary of workstations than to pulling WIPs between workstations.

2.1. CONWIP systems

Modifications have been made to the operation of the original CONWIP by Spearman et al. (1990) to cater to a specific manufacturing environment. The literature review reveals two common modifications: multiple closed loop (M-CLOSED) CONWIP and dynamic dispatch rule CONWIP. In an M-CLOSED CONWIP (Ryan, Baynat, & Choobineh, 2000), the operation is as a CONWIP system, with product
variety separated into classes based on product family and cards dedicated to each class. Aside from M-CLOSED CONWIP, dedicated cards can also be used in mixed pull systems, in which the product variety is separated into classes based on the demand for product variety. Pareto analysis is a useful tool to determine the division of classes (Grosfeld-Nir, Ronen, & Kozlovsky, 2007). The basic 80/20 principle (80% of profit held by 20% of the product variety) categorizes product variety into two classes. This principle was changed to adapt to the realities in various multi-stage multi-product environments. Pareto analysis is favored because it can integrate product families with similar processing characteristics (Scholz-Reiter, Heger, Meinecke, & Bergmann, 2012).

The classification criteria in Pareto analysis are demand, annual dollar usage, inventory cost, flow time, order size, usage regularity, and part criticality (Ramanathan, 2006). One criterion that is frequently used is demand because it is practicality, aligned with methods to address uncertainty in the demand of the product variety, and provides the basis for selecting a suitable replenishment method of raw materials (Wazed, Ahmed, & Nukman, 2009). Once product variety has been separated into classes, a suitable production control for each class is determined. For instance, in an M-CONWIP system, each class adopts a CONWIP system. In a mixed pull system, each class adopts either a kanban or a push system.

Dispatch rules can be classified as static, in which the sequence of products to be processed are determined in advance of production, or dynamic, in which the sequence uses real-time production information (Framinan, Ruiz-Usano, & Leisten, 2001). As the best dispatch rule, FIFO is widely used among other dispatch rules. Other common static dispatch rules are earliest due date (EDD) (products with the EDD are selected first) and shortest processing time (SPT) (products with the SPT at a workstation are selected first) (Panwalkar & Iskander, 1977). An example of a dynamic dispatch rule is presented in the study of Bahaji and Kuhl (2005). In this rule, products with the lowest index of a variable are selected. The variable is the sum of processing time of the product and all jobs in the queue divided by the time the product has already spent on the shop floor.

Static dispatch rules are preferred when changes in demand are deterministic, whereas dynamic dispatch rules are preferred when changes in demand are non-deterministic (Abou-Ali & Shouman, 2004). Heuristics and weighted priority indices to integrate multiple production information in dynamic dispatch rules are usually complex and difficult (Zhou & Rose, 2011). A simple solution is to apply two static dispatch rules to a queue under different circumstances (Panwalkar & Iskander, 1977). For instance, Conway, Johnson, and Maxwell (1960) studied the dollar value rule, in which products are divided into two classes: high dollar value and low dollar value. Products from the high dollar value class are selected first, and if more than one batch qualifies, selection is based on FIFO. Similarly, Scudder and Hoffmann (1987) examined a profitability rule, in which products are divided into three classes based on the profit garnered yearly. Priority is given to the class with higher profit, with selection based on FIFO within each class.

3. Model description

A generalized deterministic model of a multi-stage multi-product environment adopted in this study is presented. A finite set $J$ of $n$ orders $\{j_i\}_{i=1}^n$ is processed through a finite
set $S$ of $s$ stages $\{S_k\}_{k=1}^s$ in a fixed sequence, $S_1, S_2, \ldots, S_k, \ldots, S_s$. Each stage $S_k$ consists of a finite set $M$ of $m$ identical machines $\{M_{kl}\}_{k=1}^s$. At each stage $S_k$, each order $J_i$ is processed on a single machine $M_{kl}$. The operation of order $J_i$ on machine $M_{kl}$ consumes an uninterrupted setup time $c_{ik}$ and uninterrupted processing time $t_{ik}$.

Between stages $S_k$ and $S_{k+1}$, orders in queue $B_{k+1}$ have a specified processing sequence delineated by dispatch rule $D$. All queues of a model use an identical dispatch rule $D$. Machine $M_{kl}$ is confined by capacity constraints, which stipulate that machine $M_{kl}$ can process only one order at a time unless specified otherwise. Figure 1 illustrates the production flow in the generalized model.

Figure 1. Flow illustration of general model.

Let $b_i$ and $a_i$ be the time order $J_i$ completes processing in $S_s$ and starts processing in $S_1$, respectively. Let $N$ be the number of orders $J_i$ that have completed processing over period $T$. The objective is to determine the best production control that yields the shortest flow time ($FT_{\text{min}}$) and lowest WIP level ($WIP_{\text{min}}$), which are (Little, 1961) shown in Equations (1) and (2) respectively:

$$FT_{\text{min}} = \min(FT) = \min\left(\frac{\sum_{i=1}^n (b_i - a_i)}{N}\right)$$

$$WIP_{\text{min}} = \min(WIP) = \min\left(\frac{\sum_{i=1}^n (b_i - a_i)}{T}\right)$$

The production controls for performance comparison are benchmarked against push and CONWIP, as well as against the production control developed for this study called parallel CONWIP. Let order $J_i$ arrive at queue $B_1$, upstream of stage $S_1$. The default dispatch rule $D$ is FIFO. In the push system, order $J_i$ in queue $B_k$ is pushed to a machine $M_{kl}$ in stage $S_k$ as long as $M_{kl}$ is vacant. In the CONWIP system, order $J_i$ in queue $B_1$ is pushed to a machine $M_{kl}$ in stage $S_1$ as long as $M_{kl}$ is vacant and the number of orders in stage $S_k$ ($p_{Sk}$) and queue $B_k$ ($p_{Bk}$) for all stages and queues (except queue $B_1$) is less than $KC$. That is, as shown in Equation (3):

$$\sum_{k=1}^s p_{Sk} + \sum_{k=2}^s p_{Bk} < KC$$

In the parallel CONWIP system, the set $J$ of $n$ orders $\{J_i\}_{i=1}^n$ is split into two predefined classes, where a finite set $J_{HR}$ of $f$ orders $\{J_{HRi}\}_{i=1}^f$ is class HR, and a finite set $J_{LR}$ of $g$ orders $\{J_{LRi}\}_{i=1}^g$ is class LR. Order $J_{HRi}$ in queue $B_1$ is pushed to a machine $M_{kl}$ in stage
\( S_1 \) as long as \( M_k \) is vacant and the number of HR orders in stage \( S_k \) \((pHR_{S_k})\) and queue \( B_k \) \((pHR_{B_k})\) for all stages and queues (except queue \( B_1 \)) is less than \( KH \). That is, as shown in Equation (4):

\[
\sum_{k=1}^{s} pHR_{S_k} + \sum_{k=2}^{s} pHR_{B_k} < KH
\] (4)

Order \( J_{LRi} \) in queue \( B_1 \) is pushed to a machine \( M_{ki} \) in stage \( S_1 \) as long as \( M_{ki} \) is vacant and the number of LR orders in stage \( S_k \) \((pLR_{S_k})\) and queue \( B_k \) \((pLR_{B_k})\) for all stages and queues (except queue \( B_1 \)) is less than \( KL \). That is, as shown in Equation (5):

\[
\sum_{k=1}^{s} pLR_{S_k} + \sum_{k=2}^{s} pLR_{B_k} < KL
\] (5)

Two additional variations of parallel CONWIP are studied, each with a different dispatch rule \( D \). The first is \( HL \), in which \( J_{HRi} \) is given priority over \( J_{LRi} \), followed by \( FIFO \). The second is \( LH \), in which \( J_{LRi} \) is given priority over \( J_{HRi} \), followed by \( FIFO \). For each model, the operating parameters (i.e. \( KC \), \( KH \), and/or \( KL \)) are the state variables. One additional variable introduced is the ratio of HR orders to total orders, \( RT \). Diagrams following the examples of Krieg (2005) are presented to illustrate the operations of the production controls. Figure 2 shows the definition of the symbol used, and Figures 3–5 show the schematics of push, CONWIP, and parallel CONWIP systems, respectively.

![Figure 2. Symbols used in a diagram.](#)

| Sequence | Description |
|----------|-------------|
| 1, 3, 5  | The batch is pushed to the immediate downstream production stage |
| 2, 4, 6  | The batch is pushed to the immediate downstream buffer |

![Figure 3. Flow of material in a push system.](#)
4. Methodology

The best production control is selected based on the performance assessment using two case studies in two separate facilities. The case studies are selected based on the fulfillment of the following criteria: (a) discrete and batch production, (b) multi-stage multi-product production process in which product variety shares an identical process route, (c) no shortages in the raw material supply for the product variety in consideration, and (d) the production process is available for study. A suitable production control is usually determined together with the improvements in the production process (Smalley, 2009). Using WITNESS 2008, the models are constructed and discrete event simulation is used to compare the performances of the production controls in a given production process. In discrete event simulation, the state variable changes at discrete points in time (Law & Kelton, 2000). Discrete event simulation is used because it (Banks, Carson, Nelson, & Nicol, 2010): (a) provides a risk-free environment to test a new system, (b) handles complex systems involving uncertainty, and (c) models characteristics and movement of orders discretely.
A model is verified and validated once constructed. Comparing output with the analytical results and tracing is the verification technique, and sensitivity analysis is the validation technique (Kleijnen, 1995; Sargent, 2005). For both verification and validation, the model is first made deterministic. To compare the output with the analytical results, an equation to express the steady-state output of the model is derived mathematically (if possible) or obtained from the literature, and an intermediate trace of the output is calculated manually. The outputs from both cases are compared with the simulated output from the deterministic model. If the calculated output approximates the simulated output, the model is verified. In the sensitivity analysis, the value of a selected variable is changed incrementally, and the corresponding change in output is noted. The output in response to the variable is plotted on a graph. If the graphical relationship matches the underlying theory of the system behavior, the model is validated.

For each model, the full factorial experimental design is adopted; all combinations of the variable levels are simulated. Despite being less efficient because of the large number of experiments involved, the method provides higher accuracy for the predicted relationship between the variables and the performance measures (Montgomery, 2008). Subsequently, the warm up and run times are determined using the Welch method (Welch, 1983).

The models are ready to run when all the steps are completed. One run uses one combination of the variables, and the values of resultant performance measures are collected after each run. The resultant data for each model are analyzed using the response surface method consisting of ANOVA (to determine the significance of the variables in the performance measures) and regression (to express the performance measure as a function of the significant variable) (Montgomery, 2008) once all combinations of variables have been run. The regression equations of the production controls are depicted on graphs to select the superior system and suitable operating parameter. Figure 6 shows the flow of the methodology.

5. Case study 1

At the time of study, the facility accommodates high WIP level scattered across the shop floor. Engineers failed to trace the quantity discrepancy amounting to thousands of ringgit. Specifically, three product families (i.e. $p_1$, $p_2$, and $p_3$) sharing an identical process route (shared machines) and similar processing steps were studied. Each product family has four product varieties. Production orders arrived in batches of variable lot sizes for a single product variety. The production orders were generated from demand forecasts finalized six months prior to the due date. The finalized demand forecasts were decomposed into weekly schedules, which were sent to three production areas, namely surface mount equipment, depaneling, and back end. The push system was used for production control. The production area had two identical surface mount equipment and two identical depaneling machines. The machines in the back end were arranged in a job shop layout. Raw materials were dispatched to the shop floor according to the daily requirement.

The processing began with the labeling of printed circuit boards (PCBs). Next, the PCBs were pushed to the surface mount equipment. The processes involved were solder printing, chip placement, oven flow, and automated optical inspection. Then, the 25-in-one magazine PCBs were pushed to the depaneling machine, in which the unwanted area of the PCBs was removed through punching. Finally, 10-in-one box PCBs were pushed to the back end. The processes involved manual insert, wave soldering,
touch-up, in-circuit test, functional test, and quality assurance. The end products were packed and stored in the warehouse.

The high WIP levels located upstream of the surface mount equipment, in the depaneling machines, and at the back end were caused by several factors. First, production quantities were adjusted to account for defects, which often resulted in over-production. Second, the surface mount equipment required 30 min of online setup, with additional five hours of offline setup for components retrieval from the warehouse. Therefore, using a large lot size was essential to offset the impact of setup time on the processing time. Third, the depaneling machine handled parts from other machines. The rescheduling of orders in smaller lot sizes caused the delay in the production of magazines of PCBs for several days. Fourth, the production quantity in each area had no connection to the schedule in the upstream areas. The schedule changes in the depaneling machines were not transmitted to the back end, causing the back end to improvise based on the urgency of orders.

For the current daily output of approximately 1000 units, the management requested for (a) one dedicated depaneling machine, (b) the conversion of the back end into two dedicated U-shaped cells to allow one piece flow within each cell, (c) wave soldering to be substituted with dip soldering to reduce the production flow time, (d) removal of the
redundant functional test, (e) high-demand products to have a fixed lot size of 200 and low demand-products to follow the stipulated lot size, and (f) the retrieval of components from high-demand products in advance of the actual production and retrieval of components from low-demand products just prior to production. A suitable production control was also needed.

In the simulation models (reflecting the improved production process), three product families, namely $p_1$, $p_2$, and $p_3$, with four product varieties in each family are processed. The orders arrive every eight hours in batches with different product varieties. Each batch is composed of a single product variety. The production orders follow a fixed distribution between HR and LR ($RT$). HR batches have a lot size of 200, and LR batches have many size distributions with a mean of 100, a minimum of 50, and a maximum of 300. Production has three sequential stages, namely surface mount equipment ($SMT$), depaneling ($DEP$), and U-shaped cell ($UCELL$). $SMT$ has one machine in $DEP$ and two identical machines in $UCELL$. A batch is ready for processing once it arrives in $B_1$ and is shipped once processing in $UCELL$ is completed. LR batches have five hours of waiting time in $B_1$ to retrieve the right amount of components. Buffers $B_2$ and $B_3$ are placed between consecutive stages to decouple production. Figure 7 illustrates the production flow.

![Figure 7. Production flow of the manufacturing system in case study 1.](image)

Five production controls were tested: push and CONWIP system with FIFO dispatch rule ($PFIFO$ and $CFIFO$), parallel CONWIP system with FIFO, and HL and LH dispatch rules ($PCFIFO$, $PCHL$, and $PCLH$). In $CFIFO$, all product varieties share cards. In $PCFIFO$, $PCHL$, and $PCLH$, the division of classes between the product varieties result in two HR and two LR product varieties. This is summarized in Table 1 and

Table 1. Demand, standard deviation of demand, and coefficient of variation of the product varieties in case study 1.

| Product family | Product variety | Daily demand, $D$ | Standard deviation of demand, $\sigma_D$ | Coefficient of variation, $C_v = \frac{\sigma_D}{D}$ |
|---------------|----------------|------------------|----------------------------------------|--------------------------------|
| $p_1$         | $A$            | 105              | 30                                     | .29                           |
|               | $B$            | 84               | 12                                     | .14                           |
|               | $C$            | 15               | 10                                     | .67                           |
|               | $D$            | 8                | 7                                      | .88                           |
| $p_2$         | $E$            | 391              | 94                                     | .24                           |
|               | $F$            | 296              | 20                                     | .07                           |
|               | $G$            | 36               | 21                                     | .58                           |
|               | $H$            | 27               | 16                                     | .59                           |
| $p_3$         | $I$            | 257              | 37                                     | .14                           |
|               | $J$            | 99               | 8                                      | .08                           |
|               | $K$            | 28               | 19                                     | .68                           |
|               | $L$            | 20               | 17                                     | .85                           |
Figure 8 below. In theory, products with higher order quantity have a more stable demand, while products with lower order quantity are more sporadic. This is consistent with Figure 8, where the left of the dotted line is of products with low demand and high variability in quantity (LR), while the right is of products with high demand and low variability in quantity (HR).

The warm up and run times were 518,400 and 1,555,200 s, respectively. Tables 2 and 3 show the parameter and variable values of the models of case study 1. The values of RT were selected based on the records of production planning (i.e. minimum, average, and maximum demand of HR orders from total orders). The values of KC, KH, and KL were selected to show visible changes in the performance measures. This selection was done because the improved production process is significantly different from the existing process.

Table 2. Parameter values of case study 1 models.

| Machine Product family | SMT | DEP | UCELL |
|------------------------|-----|-----|-------|
|                        | p₁  | p₂  | p₃    |
|                        | p₁  | p₂  | p₃    |
| Cycle time per product (s) | 54  | 47  | 125  |
| Setup time per batch (s) | 1800 | 1800 | 3600 |

Table 3. Discrete values of variables for case study 1 models.

| Production control | Variable |
|--------------------|----------|
| PFIFO              | RT=.5, .7, .9, KC=KH=KL=– |
| CFIFO              | RT=–, KC=4, 8, 12, KH=KL=– |
| PCFIFO             | RT=–, KC=2, 4, 6, KH=KL=– |
| PCHL               | RT=–, KC=2, 4, 6, KH=KL=2, 4, 6 |
| PCLH               | RT=–, KC=2, 4, 6, KH=KL=2, 4, 6 |
The results of the ANOVAs are first presented. For PFIFO, RT has a significant effect on FT and WIP. For CFIFO, RT has a significant effect on WIP, and KC has a significant effect on FT. For PCFIFO, all factors and interactions have significant effects on FT and WIP. For PCHL, only KH and RT and KH interaction have significant effects on FT and WIP. Similarly for PCLH, only KL and RT and KL interaction have significant effects on FT and WIP.

All five systems cannot be plotted in a single graph because they have no common variable with a significant effect. This problem is resolved by plotting two graphs for each performance measure. The first graph uses RT as the variable for PFIFO and PCFIFO. The second uses KC as the variable for CFIFO, PCFIFO, PCHL, and PCLH. KC is equal to the sum of KH and KL. The formation of a new variable (KC) from two individual variables (KH + KL) is valid as long as the new variable has a significant effect on the performance measure. The graphs are shown in Figures 9–12.

Figure 9. FT against RT of case study 1.

Figure 10. FT against KC of case study 1.
PCFIFO has a lower FT than PFIFO but at above a certain cutoff point only \((RT = .55)\) (Figure 9). However, PCFIFO has the advantage of a steady and, thus, more predictable FT. Compared with PCHL and PCLH (Figure 10), PCFIFO has higher FT. However, the advantage of PCFIFO lies in the smaller variance of FT. The same advantage is also observed for WIP (Figures 11 and 12). Therefore, PCFIFO is most suitable based on this assessment, with 8 KC (4 KH and 4 KL) at \(FT_{\text{min}}\) and \(WIP_{\text{min}}\).

6. Case study 2

The facility that manufactures composite sub-assemblies for a well-known commercial aircraft expressed interest in using the pull system for production control. At the time of study, the WIP level on the shop floor was at an all-time high. However, to justify a more pervasive usage, a trial run on a single product family was required. Minimal adjustments were also made to processes shared by other product families.

The product family selected was \(p\), which contained four product varieties (i.e. \(p_1\), \(p_2\), \(p_3\), and \(p_4\)). The weekly demand for \(p\) was 10 sets. Each set was composed of two non-identical panels. Demand forecasts were received one year in advance and finalized...
two months before the scheduled shipment. Once finalized, the schedule was sent to the three main production areas, such as sub-assembly, paint shop, and final assembly. In sub- and final assemblies, the workstations were arranged in a dedicated cell layout, and a generalized machine layout in the paint shop showed the shared equipment among various product families. Raw panels were delivered from inbound to sub-assembly based on reorder quantity (ROQ), and the remaining components were delivered based on reorder point (ROP). In ROQ, replenishment occurs at fixed intervals, whereas in ROP, replenishment occurs when inventory hits a minimum.

A push system was used for production control. In the sub-assembly, a set underwent five processes, namely, in-jig, out-jig, curing (24 h), out-jig wet, curing (1 h), and touch-up. These processes involved drilling holes in the panels, inserting fittings into the holes, fixing the fittings with sealant, and air curing of the panels. Next, the set was pushed to the paint shop and underwent five processes, namely part preparation 1, primer coating, part preparation 2, top coating, and touch-up. These processes involved filling and gritting to level the surface and paint coating of the panel on the defined areas of the surface. Finally, the set was pushed to the final assembly and underwent assembly of the remaining components prior to an additional six hours of curing, followed by touch-up. The end product was packed and stored in outbound.

High WIP levels at the upstream of the sub-assembly, paint shop, and final assembly were attributed to two factors. First, the use of ROQ led to the delivery of raw panels to the sub-assembly irrespective of their needs. Second, the batch production at the paint shop had to wait for the right quantity of sets before processing began. Consequently, a set was frequently retained for several days at the final assembly as the daily capacity of the shop could only process two sets from the batch. The management requested for (a) the replenishment of raw panels to be changed from ROQ to ROP to ensure that the delivery of raw panels matched the requirements of the sub-assembly and (b) the determination of the minimum quantity of WIPs in the sub-assembly, paint shop, and final assembly to establish the card count.

One method to set the minimum WIP level is known as the ‘greedy’ procedure (Federgruen & Groenevelt, 1986), in which each set is attached with a card and decreased progressively by a small amount. Table 4 summarizes the results of the ‘greedy’ procedure. $B_1$, $B_2$, $B_3$, and $B_4$ are placed at the immediate upstream of out-jig wet, part preparation 1, part preparation 2, and final assembly, respectively. Table 4 shows that 14 ($2 + 4 + 4 + 4 = 14$) cards were required. Additional four cards were introduced into the system: two for a daily delivery of two sets of raw panels and two for a daily output of two sets. Therefore, 18 ($14 + 2 + 2 = 18$) cards were required.

| Week | $B_1$ | $B_2$ | $B_3$ | $B_4$ |
|------|-------|-------|-------|------|
| 1    | 7     | 3     | 5     | 4    |
| 2    | 6     | 4     | 4     | 4    |
| 3    | 5     | 5     | 4     | 4    |
| 4    | 4     | 4     | 4     | 4    |
| 5    | 3     | 4     | 4     | 4    |
| 6    | 2     | 4     | 4     | 4    |

Table 4. Results of ‘greedy’ procedure.
In the simulation models (integrating the improvements), one product family, \( p \), consisting of four product varieties, \( p_1, p_2, p_3, \) and \( p_4 \), are processed. \( p_1 \) is HR, and \( p_2, p_3, \) and \( p_4 \) are LR. Orders arrive in sets of two every eight hours. Each set comprises two non-identical panels from a single product variety. The production orders follow a fixed distribution between HR and LR (RT). Production proceeds in five sequential stages, namely \( IO \) (in-jig and out-jig), \( OT \) (out-jig wet, 1-h curing and touch-up), \( PP \) (part preparation 1 and primer coating), \( PT \) (part preparation 2, top coating and touch-up), and \( FA \) (final assembly). In each stage, the cycle time is independent of the product variety being processed. A set is ready for processing once it arrives in \( B_0 \). \( B_1, B_2, B_3, \) and \( B_4 \) are placed between stages to decouple the production. \( B_1 \) and \( B_5 \) represent curing operations with curing times of 24 and 6.3 h, respectively. A set is shipped once curing is completed in \( B_5 \). In \( PP \), a minimum of four sets is required before processing can begin. In \( PT \), a minimum of four sets of HR or two sets of LR are required before processing can begin. Figure 13 illustrates the production flow.

![Figure 13. Production flow of the manufacturing system in case study 2.](image)

The five production controls are tested, and the selection of RT follows that of case study 1. The division between classes HR and LR in \( PCFIFO, PCHL \) and \( PCLH \) is summarized in Table 5 and Figure 14. In Figure 14, the left of the dotted line shows product varieties of low demand and high variability in quantity (\( B, C, \) and \( D \) overlap), while the right shows the product variety with high demand and low variability in quantity (\( A \)).

| Product variety | Monthly demand, \( D \) | Standard deviation of demand, \( \sigma_D \) | Coefficient of variation, \( C_v = \sigma_D/D \) |
|-----------------|--------------------------|------------------------------------------|---------------------------------|
| \( A \)         | 36                       | 9                                        | .25                             |
| \( B \)         | 1                        | 1                                        | 1.00                            |
| \( C \)         | 1                        | 1                                        | 1.00                            |
| \( D \)         | 1                        | 1                                        | 1.00                            |

The warm up and run times are 28 and 140 h, respectively. Tables 6 and 7 show the parameter and variable values of case study 2 models.

The results of the ANOVAs are first presented. For \( CFIFO \) and \( PFIFO \), only RT has a significant effect on \( FT \) and \( WIP \). In \( PCFIFO, PCHL \), and \( PCLH, RT, KL \), and \( RT \) and \( KL \) interaction have significant effects on \( FT \) and \( WIP \). The graphs \( FT \) and \( WIP \) for the five production controls are shown in Figures 15 and 16, respectively.
All systems (except PCHL) have the same FT (Figure 15). At RT of .9, all systems have identical WIP (Figure 16). When RT is .7, PFIFO, CFIFO, and PCLH have identical WIP. Therefore, PFIFO, CFIFO, and PCLH are the most suitable.

7. Discussion

The ANOVAs of both case studies reveal one distinct observation in PCFIFO, PCHL, and PCLH that are absent in PFIFO and CFIFO: the significance of KH/KL interactions with RT. The discrete RT values represent a fixed ratio between HR orders and total orders. A change from one discrete value to another discrete value represents a change
in the quantity of HR and LR orders. Changes in HR and LR orders are, in turn, a response to changes in HR and LR demands.

Demand forecasts are updated periodically in the MRP II. This process works well for MTS products, in which the forecast horizon is far ahead from the lead time. For MTO products, in which production is initiated by demand, the forecast must be finalized long enough to cover the lead time. Manufacturing organizations will always attempt to maintain a fixed cost in production and partially reduce these costs during slow economic conditions (Swamidass, 2000). Once the organization establishes a lot size that minimizes the total cost, the tendency is to retain that lot size (Cimorelli, 2013). Therefore, the lot size is usually fixed for the short term. For the push system, demand forecasts are translated to production orders of fixed lot sizes and released to the production at scheduled times (Bonvik, Couch, & Gershwin, 1997). For pull systems, demand forecasts are translated to production orders of a predefined card quantity (each bearing a fixed lot size), with the production regulated by the cards (Karmarkar, 1991).

The ANOVAs of PFIFO for both case studies reveal that RT is the only factor that has a significant effect on FT and WIP because RT is the only variable present. Therefore, all changes in performance measures are in response to changes in RT. In effect,
Table 8. Summary of findings for case studies 1 and 2.

| Performance measure | Ranking | *Case study 1 | **Case study 2 | Performance measure | Ranking | *Case study 1 | ***Case study 2 |
|---------------------|---------|---------------|----------------|---------------------|---------|---------------|----------------|
| Increasing FT (less preferred) | 1       | PCHL          | PCLH           | Increasing WIP (less preferred) | 1       | PCHL          | PCLH           |
|                     | 2       | PCLH          | PFIFO (benchmark) |                   | 2       | PCLH          | PFIFO (benchmark) |
|                     | 3       | PCFIFO        | CFIFO (benchmark) |                   | 3       | PCFIFO        | CFIFO (benchmark) |
|                     | 4       | CFIFO (benchmark) | PCFIFO        |                   | 4       | PFIFO (benchmark) | PCFIFO        |
|                     | 5       | PFIFO (benchmark) | PCHL          |                   | 5       | CFIFO (benchmark) | PCHL          |

*FT and WIP of PCHL and PCLH are lower, but with high variance.
**PCLH, PFIFO, CFIFO, and PCFIFO have overlapping FT.
***PCLH, PFIFO, and CFIFO have overlapping WIP.
the arrival of HR and LR orders with predefined lot sizes already reflects the changes in HR and LR demands.

The CFIFO findings in both case studies differ. In case study 1, the significance of KC indicates that card count responds to demand changes. However, Marek, Elkins, and Smith (2001) proved that the setting of card count in CONWIP systems is based on the total demand. As CONWIP systems only respond to changes in total demand, the demand variation is equal across all product families (Jodlbauer & Huber, 2008). In fact, the demand variation is not equal based on the principle of Pareto analysis, and thus separate classes are required. Therefore, CONWIP systems resort to other methods, such as M-CLOSED CONWIP, which results in increased WIP level and holding costs. In case study 2, only RT has a significant effect in demonstrating that the setting of KC is high. This behavior is similar to that in PFIFO, indicating that CFIFO has approximated a push system as KC does very little to limit the WIP level. Nevertheless, as a generalization for CFIFO, the absence of KC and RT interaction indicates that the card count does not respond to changes in HR and LR demands.

The ANOVAs reveal that the KH and RT interaction and/or KL and RT interaction is only significant in parallel CONWIP systems. In other words, only the parallel CONWIP system responds to changes in HR and LR demands. Therefore, a selected card count serves as a suitable starting point for more quantities depending on RT.

The graphical analysis from both case studies demonstrates the different production control preferences. In case study 1, PCFIFO is preferred because of (a) low FT and WIP compared with PFIFO and CFIFO and (b) low variance in FT and WIP compared with PFIFO, CFIFO, PCHL, and PCLH. However, in case study 2, PFIFO, CFIFO, and PCLH are preferred. The batch production in PT (four sets of HR or two sets of LR are required before processing can begin) is a disadvantage to PCFIFO and PCHL because the arrival of the required class of sets at PT is dependent on the withdrawal of completed sets from the same class. In CFIFO, the arrival of the required class of sets at PT is dependent on the withdrawal of completed sets from any class. In PFIFO, the arrival of the required class of sets at PT is dependent on the availability of the raw panels. In PCLH, when HR demand is low (RT = .7), the quantity of HR remains relatively high compared with that of LR. However, one concern is the accumulation of the two LR sets upstream of PT. With the LH dispatch rule, this problem is solved as LR is given priority over HR. In theory, the performance improves when products with low demands are prioritized (Miller, Ginsberg, & Maxwell, 1974).

In the absence of batch production, PCFIFO performs better than PFIFO, CFIFO, PCHL, and PCLH for FT and WIP (case study 1). In the presence of batch production, PCLH, PFIFO, and CFIFO perform better than PCFIFO and PCHL for FT and WIP (case study 2). Table 8 summarizes the findings from both case studies, with the choice production control(s) in bold.

8. Conclusion

The paper presents a production control in a multi-stage multi-product environment with the objective of reducing the flow time and WIP level. The production control is targeted toward environments where product mixes have different demands and the fulfillment of the demands while maintaining a low WIP level and short flow time is required. The proposed production control was examined in two case studies that are significantly different in terms of the application environment. In the first case study, parallel CONWIP system was benchmarked against the push and CONWIP systems in
an electronics assembly facility through discrete event simulation. The models reflected the changes in the production process as well as the induction of the production control. The results reveal the superiority of the parallel CONWIP with FIFO dispatch rule for WIP level and flow time. In the second case study, an aero structure manufacturing facility considered the minimal changes to the production process and batch production at the production stage. The results reveal the superiority of push, CONWIP, and parallel CONWIP with LH dispatch rule for WIP level and flow time.

The simulation results show that the parallel CONWIP system yields good performance. Nevertheless, one notable advantage is the ability of card count to naturally respond to changes in the demand of product classes. Several aspects can be considered for future studies. First, given that parallel CONWIP systems respond to demand changes, a framework of dynamic card controlling can provide insights into their reactive nature. Second, the consideration of defects in the analysis will be beneficial. Opportunities lie in the rework strategy employed and the quantity of defects. Third, the integration of parallel CONWIP systems with dynamic dispatch rules that incorporate other aspects of production (e.g. longest remaining processing time) can be examined. Fourth, studying the implementation of parallel CONWIP systems aligned with lean practices is useful.

Funding
The research presented is cosponsored by Universiti Sains Malaysia short-term research [grant number 60311042]; APEX [grant number 910345].

References
Abou-Ali, M., & Shouman, M. (2004). Effect of dynamic and static dispatching strategies on dynamically planned and unplanned FMS. Journal of Materials Processing Technology, 148, 132–138.
Addison, D. M. (2003). Productivity growth and product variety: Gains from imitation and education. World Bank Policy Research Paper 3023, Washington, DC.
Alford, D., Sackett, P., & Nelder, G. (2000). Mass customisation – An automotive perspective. International Journal of Production Economics, 65, 99–110.
Araya, J. M. (2012). Value stream mapping adapted to high-mix, low-volume manufacturing environments (PhD thesis). KTH Industrial Engineering and Management.
Bahaji, N., & Kuhl, M. E. (2005). A simulation study of new multi-objective composite dispatching rules, CONWIP, and push lot release in semiconductor fabrication. Journal of Production Research, 46, 3801–3824.
Banks, J., Carson, J. S., Nelson, B. L., & Nicol, D. M. (2010). Discrete-event system simulation. Upper Saddle River, NJ: Pearson Education.
Bar, S. N. (2006). Business expansion and lean transformation for helicopter blade shop (PhD thesis). Massachusetts Institute of Technology.
Barba-Gutiérrez, Y., & Adenso-Díaz, B. (2009). Reverse MRP under uncertain and imprecise demand. The International Journal of Advanced Manufacturing Technology, 40, 413–424.
Bils, M., & Klenow, P. J. (2001). The acceleration in variety growth. American Economic Review, 91, 274–280.
Bonvik, A. M., Couch, C. E., & Gershwin, S. B. (1997). A comparison of production-line control mechanisms. International Journal of Production Research, 35, 789–804.
Cimorrelli, S. (2013). Kanban for the supply chain: Fundamental practices for manufacturing management. New York, NY: Productivity Press.
Conway, R. W., Johnson, B. M., & Maxwell, W. L. (1960). An experimental investigation of priority dispatching. Journal of Industrial Engineering, 11, 221–229.
Denton, B., Gupta, D., & Jawahir, K. (2003). Managing increasing product variety at integrated steel mills. Interfaces, 33, 41–53.
El-Khouly, I. A., El-Kilany, K. S., & El-Sayed, A. E. (2009). Modelling and simulation of re-entrant flow shop scheduling: An application in semiconductor manufacturing (pp. 211–216). Proceedings of International Conference on Computers & Industrial Engineering, Troyes.

Federgruen, A., & Groenevelt, H. (1986). The greedy procedure for resource allocation problems: Necessary and sufficient conditions for optimality. *Operations Research, 34*, 909–918.

Framinan, J. M., González, P. L., & Ruiz-Usano, R. (2003). The CONWIP production control system: Review and research issues. *Production Planning & Control, 14*, 255–265.

Framinan, J. M., Ruiz-Usano, R., & Leisten, R. (2001). Sequencing CONWIP flow-shops: Analysis and heuristics. *International Journal of Production Research, 39*, 2735–2749.

Gao, G., & Hitt, L. M. (2004). Information technology and product variety: Evidence from panel data. *Information Technology, 12*, 365–378.

Gates, M. D. (2004). *Lean manufacturing system design and value stream management in a high-mix, low-volume environment* (Master thesis). Massachusetts Institute of Technology.

Germs, R., & Riezebos, J. (2010). Workload balancing capability of pull systems in MTO production. *International Journal of Production Research, 48*, 2345–2360.

Grosfeld-Nir, A., Ronen, B., & Kozlovsky, N. (2007). The Pareto managerial principle: When does it apply? *International Journal of Production Research, 45*, 2317–2325.

Gurgur, C. Z., & Altıok, T. (2008). Decentralized multi-product multi-stage systems with backorders. *IIE Transactions, 40*, 238–251.

Harrison, D. K., & Petty, D. J. (2002). *Systems for planning and control in manufacturing*. Woburn, MA: Butterworth-Heinemann.

Hopp, W. J., & Roof, M. L. (1998). Setting WIP levels with statistical throughput control (STC) in CONWIP production lines. *International Journal of Production Research, 36*, 867–882.

Hopp, W. J., & Spearman, M. L. (2001). *Factory physics: The foundations of manufacturing management*. New York, NY: McGraw-Hill.

Hopp, W. J., & Spearman, M. L. (2004). To pull or not to pull: What is the question? *Manufacturing & Service Operations Management, 6*, 133–148.

Horbal, R., Kagan, R., & Koch, T. (2008). Implementing lean manufacturing in high-mix production environment. In T. Koch (Ed.), *Lean business systems and beyond, IFIP International Federation for Information Processing* (Vol. 257, pp. 257–267). Boston, MA: Springer.

Jodlbauer, H., & Huber, A. (2008). Service-level performance of MRP, kanban, CONWIP and DBR due to parameter stability and environmental robustness. *International Journal of Production Research, 46*, 2179–2195.

Kanet, J. J., & Stoßlein, M. (2010). Integrating production planning and control: Towards a simple model for capacitated ERP. *Production Planning & Control, 21*, 286–300.

Karmarkar, U. S. (1991). Push, pull and hybrid control schemes. *Tijdschrift voor Economie en Management, 26*, 345–363.

Katō, I., & Smalley, A. (2011). *Toyota kaizen methods: Six steps to improvement*. New York, NY: Productivity Press.

Khojasteh-Ghamari, Y. (2009). A performance comparison between kanban and CONWIP controlled assembly systems. *Journal of Intelligent Manufacturing, 20*, 751–760.

Kleijnen, J. P. (1995). Verification and validation of simulation models. *European Journal of Operational Research, 82*, 145–162.

Krieg, G. N. (2005). *Kanban-controlled manufacturing systems*. Berlin: Springer-Verlag.

Krieg, G. N., & Kuhn, H. (2008). Performance evaluation of two-stage multi-product kanban systems. *IIE Transactions, 40*, 265–283.

Krishnamurthy, A., & Suri, R. (2009). Planning and implementing POLCA: A card-based control system for high variety or custom engineered products. *Production Planning & Control, 20*, 596–610.

Kumar, C. S., & Panneerselvam, R. (2007). Literature review of JIT-KANBAN system. *The International Journal of Advanced Manufacturing Technology, 32*, 393–408.

Lancaster, K. J. (1990). The economics of product variety: A survey. *Marketing Science, 9*, 189–206.

Lavoie, P., Gharbi, A., & Kenné, J. P. (2010). A comparative study of pull control mechanisms for unreliable homogenous transfer lines. *International Journal of Production Economics, 124*, 241–251.

Law, A. M., & Kelton, W. D. (2000). *Simulation modelling and analysis*. New York, NY: McGraw-Hill.
Lee, C., & Chen, C. (1997). A dispatching scheme involving move control and weighted due date for wafer foundries. *IEE Transactions on Components, Packaging, and Manufacturing Technology, 20*, 268–277.

Lee, L. C. (1989). A comparative study of the push and pull production systems. *International Journal of Operations & Production Management, 9*, 5–18.

Li, B. (2009). *Making a high-mix make-to-order production system lean* (PhD thesis). Massachusetts Institute of Technology.

Little, J. D. C. (1961). A proof for the queuing formula: $L = \lambda W$. *Operations Research, 9*, 383–387.

Lyons, A. C., Mondragon, A. E. C., Piller, F., & Poler, R. (2012). *Customer-driven supply chains: From glass pipelines to open innovation networks*. London: Springer Verlag.

MacDuffie, J. P., Sethuraman, K., & Fisher, M. L. (1996). Product variety and manufacturing performance: Evidence from the international automotive assembly plant study. *Management Science, 42*, 350–369.

Mahadevan, B. (2010). *Operations management: Theory and practice*. New Delhi: Pearson Education India.

Marek, R. P., Elkins, D. A., & Smith, D. R. (2001). *Manufacturing controls: Understanding the fundamentals of kanban and CONWIP pull systems using simulation* (pp. 921–929). Proceedings of the 33rd Winter Simulation Conference, Phoenix, AZ.

McKay, K. N., & Wiers, V. C. (2004). *Practical production control: A survival guide for planners and schedulers*. Boca Raton, FL: J. Ross.

Miller, L. W., Ginsberg, A. S., & Maxwell, W. L. (1974). *An experimental investigation of priority dispatching in aircraft maintenance, using a simplified model*. Santa Monica, CA: Rand Corporation.

Mohammaditabar, D., Ghodsypour, S. H., & O’Brien, C. (2012). Inventory control system design by integrating inventory classification and policy selection. *International Journal of Production Economics, 140*, 655–659.

Mönch, L. (2005). *Simulation-based assessment of order release strategies for a distributed shifting bottleneck heuristic* (pp. 2186–2193). Proceedings of the 37th Winter Simulation Conference, Orlando, FL.

Monden, Y. (1983). *Toyota production systems*. Atlanta, GA: Industrial Engineering and Management Press.

Montgomery, D. C. (2008). *Design and analysis of experiments*. New York, NY: Wiley.

Mula, J., & Poler, R. (2010). Fuzzy material requirement planning. In *Production engineering and management under fuzziness* (pp. 39–57). Heidelberg-Berlin: Springer.

Panwalkar, S. S., & Iskander, W. (1977). A survey of scheduling rules. *Operations Research, 25*, 45–61.

Parvin, H., Van Oyen, M. P., Pandelis, D. G., Williams, D. P., & Lee, J. (2012). Fixed task zone chaining: Worker coordination and zone design for inexpensive cross-training in serial CONWIP lines. *IIE Transactions, 44*, 894–914.

Ramanathan, R. (2006). ABC inventory classification with multiple-criteria using weighted linear optimization. *Computers & Operations Research, 33*, 695–700.

Ramdas, K. (2003). Managing product variety: An integrative review and research directions. *Production and Operations Management, 12*, 79–101.

Riezebos, J. (2006). *Polca simulation of a unidirectional flow system*. Proceedings of the Third International Conference on Group Technology/Cellular Manufacturing, Groningen, Netherlands.

Robinson, E. P., Sahin, F., & Gao, L. (2005). The impact of e-replenishment strategy on make-to-order supply chain performance. *Decision Sciences, 36*, 33–64.

Rose, O. (2001). *CONWIP-like lot release for a wafer fabrication facility with dynamic load changes* (pp. 41–45). Proceedings of the 2001 International Conference on Semiconductor Operational Modeling and Simulation, Seattle, WA.

Ryan, S. M., Baynat, B., & Choobineh, F. (2000). Determining inventory levels in a CONWIP controlled job shop. *IIE Transactions, 32*, 105–114.

Sargent, R. G. (2005). *Verification and validation of simulation models* (pp. 130–143). Proceedings of the 37th Winter Simulation Conference, Orlando, FL.
Saurin, T. A., Marodin, G. A., & Ribeiro, J. L. D. (2011). A framework for assessing the use of lean production practices in manufacturing cells. *International Journal of Production Research, 49*, 3211–3230.

Scholz-Reiter, B., Heger, J., Meinecke, C., & Bergmann, J. (2012). Integration of demand forecasts in ABC-XYZ analysis: Practical investigation at an industrial company. *International Journal of Productivity and Performance Management, 61*, 445–451.

Schutt, J. H. (2004). *Directing the flow of product: A guide to improving supply chain planning*. Ft. Lauderdale, FL: J. Ross.

Scudder, G. D., & Hoffmann, T. R. (1987). The use of cost-based priorities in random and flow shops. *Journal of Operations Management, 7*, 217–232.

Skelley, K. B. (2009). *Lean implementation to improve scheduling for a multi-cell manufacturing facility* (PhD thesis). University of Wisconsin.

Slomp, J., Bokhorst, J. A., & Germs, R. (2009). A lean production control system for high-variety/low-volume environments: A case study implementation. *Production Planning & Control, 20*, 586–595.

Smalley, A. (2009). *Creating level pull*. Cambridge, MA: Lean Enterprise Institute.

Spearman, M. L., Woodruff, D. L., & Hopp, W. J. (1990). CONWIP: A pull alternative to kanban. *International Journal of Production Research, 28*, 879–894.

Spearman, M. L., & Zazanis, M. A. (1992). Push and pull production systems: Issues and comparisons. *Operations Research, 40*, 521–532.

Swamidass, P. (2000). *Encyclopedia of production and manufacturing management*. Boston, MA: Kluwer Academic Publishers.

Vandaele, N., Nieuwenhuyse, I. V., Claerhout, D., & Cremmery, R. (2008). Load-based POLCA: An integrated material control system for multiproduct, multimachine job shops. *Manufacturing & Service Operations Management, 10*, 181–197.

Villa, A. (1989). Decision architectures for production planning in multi-stage multi-product manufacturing systems. *Annals of Operations Research, 17*, 51–67.

Wang, J. X. (2010). *Lean manufacturing: Business bottom-line based*. Boca Raton, FL: Taylor & Francis.

Wazed, M., Ahmed, S., & Nukman, Y. (2009). Uncertainty factors in real manufacturing environment. *Australian Journal of Basic and Applied Sciences, 3*, 342–351.

Welch, P. (1983). The statistical analysis of simulation results. In S. Lavenberg (Ed.), *The computer modeling handbook* (pp. 268–328). New York, NY: Academic Press.

Yang, T., Hsieh, C. H., & Cheng, B. Y. (2011). Lean-pull strategy in a re-entrant manufacturing environment: A pilot study for TFT-LCD array manufacturing. *International Journal of Production Research, 49*, 1511–1529.

Zhou, Z., & Rose, O. (2011). *A composite rule combining due date control and WIP balance in a wafer fab* (pp. 2080–2287). Proceedings of the Winter Simulation Conference, Phoenix, AZ.