

POLC-A: An Assessment of POLCA's Authorization Element

Matthias Thürer (corresponding author), Nuno O. Fernandes, Mark Stevenson, Cristovao Silva,
and Silvio Carmo-Silva

Name: Professor Matthias Thürer
Institution: Jinan University
Address: Institute of Physical Internet
School of Electrical and Information Engineering
Jinan University (Zhuhai Campus)
519070, Zhuhai, PR China
E-mail: matthiasthurer@workloadcontrol.com

Name: Professor Nuno O. Fernandes
Institution: Instituto Politécnico de Castelo Branco
Address: Av. do Empresário
6000-767 Castelo Branco - Portugal
E-mail: nogf@ipcb.pt

Name: Professor Mark Stevenson
Institution: Lancaster University
Address: Department of Management Science
Lancaster University Management School
Lancaster University
LA1 4YX, U.K.
E-mail: m.stevenson@lancaster.ac.uk
Tel: 00 44 1524 593847

Name: Professor Cristóvão Silva
Institution: CEMUC - University of Coimbra
Address: Mechanical Engineering Department
Polo II Pinhal de Marrocos
3030 Coimbra – Portugal
Email: cristovao.silva@dem.uc.pt

Name: Professor Silvio Carmo-Silva
Institution: University of Minho
Address: ALGORITMI Research Unit
Campus de Gualtar
4710-057 Braga – Portugal
Email: scarmo@dps.uminho.pt

Keywords: *Order Release; POLCA; Quick Response Manufacturing; Dispatching; Card-based Control.*

POLC-A: An Assessment of POLCA's Authorization Element

Abstract

POLCA (i.e. Paired-cell Overlapping Loops of Cards with Authorization) is a card-based production control approach developed to support the adoption of Quick Response Manufacturing. POLCA's control mechanism is unique since it combines a card-based element (the paired cell overlapping loops of cards) with a higher-level Material Requirements Planning (HL/MRP) system for release authorization. POLCA has been applied in practice and evaluated in research, but the loops of cards element (POLC) has been adopted without the authorization element (A). In response, we use simulation to evaluate the effect of POLCA's authorization element. We show that this element has a direct detrimental effect on percentage tardy and mean tardiness performance. While the literature argues that the authorization element should be an integral part of POLCA, our results suggest the opposite. This has important implications for research and practice. Instead of using POLCA with its authorization element, it is preferable to combine POLC – the card-based element – with a shop floor dispatching rule.

Keywords: *Order Release; POLCA; Quick Response Manufacturing; Dispatching; Card-based Control.*

1. Introduction

POLCA (i.e. Paired-cell Overlapping Loops of Cards with Authorization) is a production planning and control system that combines a card-based element (the Paired-cell Overlapping Loops of Cards element: “POLC”) with a higher-level Material Requirements Planning (HL/MRP) system for release Authorization (“A”). In other words, the use of a POLCA card to signal available capacity at the next station in the routing of an order is combined with a second, authorization element. The focus of this study is on assessing, for the first time, the impact of the authorization element on performance.

POLCA is an input control system that controls the flow of work onto the shop floor. Suri (1998) was the first to present POLCA as a production planning and control approach to support the adoption of his Quick Response Manufacturing (QRM) philosophy, or the pursuit of time-based competition (Stalk, 1990). POLCA has been argued to be an alternative to *kanban* systems specifically for companies that produce a high variety of products on a to-order basis (e.g. Riezebos, 2010; Gonzalez *et al.*, 2012). QRM and POLCA are robust systems that can aid in situations where variety is high and lean manufacturing cannot be applied (Stump & Badurdeen, 2012).

Yet although a number of studies in the extant literature report on implementations of POLCA in practice (e.g. Vandaele *et al.*, 2008; Krishnamurthy & Suri, 2009; Riezebos, 2010) and/or on assessments of POLCA’s performance using simulation (e.g. Lödding *et al.*, 2003; Germs & Riezebos, 2010; Harrod & Kanet, 2013; Severino & Godinho Filho, 2017), an integral aspect of POLCA appears to have been widely neglected. To the best of our knowledge, prior studies have neither implemented nor assessed POLCA’s authorization element. The focus has instead been on POLCA’s card-based element (i.e. POLC), which has been supported by a priority dispatching rule on the shop floor for choosing between queuing orders. POLCA’s authorization element (A) is unique since it prohibits a station from working on a job if an earliest release date calculated by the HL/MRP system has not been reached. This element is very different from the use of a dispatching rule, as typically applied in make-to-order contexts. While both seek to prioritize the jobs that are queuing in front of a station, the HL/MRP authorization element restricts the set of eligible jobs.

Although the POLCA literature has emphasized the integral importance of POLCA’s authorization element, its impact on performance remains largely unknown. Authors such as Vandaele *et al.* (2008) and Suri (2010) have argued that it avoids earliness, but this claim needs to be further unpacked. For example, delaying production will reduce earliness, but this is typically

only achieved at the expense of a higher percentage of tardy orders and a higher mean tardiness. It is the latter performance measures that are arguably most important in a make-to-order context; but the POLCA literature provides no insight into the impact of the authorization element on tardiness performance. Although Riezebos (2010) argued that the authorization element might improve tardiness performance, this claim has not been verified.

The above represents a major shortcoming in the POLCA literature. Indeed, if we look at the related literature on premature idleness, i.e. where a station is starving although there is work waiting to be processed, there is reason to fundamentally question the use of an authorization element for withholding work (Kanet, 1988; Land & Gaalman, 1998; Thürer *et al.*, 2012; Fernandes *et al.*, 2016). POLCA's authorization element is likely to be very active in periods when jobs arrive ahead of their ERD and inactive when jobs are lagging behind. The former situation can typically be associated with periods of low load while the latter situation can typically be associated with periods of high load (Land *et al.*, 2015). As a consequence, POLCA's authorization element may shift work from low load periods to already congested high load periods, negatively affecting due date performance. But research is required to confirm whether this effect holds in the context of POLCA, thereby supporting researchers and practitioners in deciding whether to implement POLCA in its entirety. In response, this study uses simulation to assess the performance impact of POLCA's authorization element. Although the authorization element is widely discussed as an integral part of POLCA, this is to the best of our knowledge the first time that its impact is actually assessed.

The remainder of this paper is structured as follows. In Section 2, we review the mechanisms underlying POLCA systems. The simulation model used to evaluate performance is then described in Section 3 before the results are presented, discussed, and analyzed in Section 4. Finally, conclusions are drawn in Section 5, where managerial implications and future research directions are also outlined.

2. Background – The POLCA Production Planning and Control System

The aim of this section is to outline the POLCA system in order to provide the insight into its underlying mechanisms that is required for our argumentation. We do not present a systematic literature review since this has been provided within the work of Riezebos (2010) and Severino & Godinho Filho (2017). Rather, we discuss the literature through the lens of POLCA's authorization element in Section 2.3. But first, POLCA's card-based element and POLCA's HL/MRP based authorization element are described in Section 2.1 and Section 2.2, respectively.

2.1 Mechanisms Underlying a POLCA System – Paired-Cell Overlapping Loop of Cards

POLCA links together the different stations in the routings of orders using card loops. We use the term ‘station’ throughout the paper, but the term ‘cell’ could equally have been applied. In fact, the POLCA literature often refers to the use of cells. But this is just a question of the level of analysis. For the functioning of the system, it is irrelevant whether stations, lines, cells, companies, etc. are controlled. For example, *kanban* is typically associated with stations yet, more than 30 years ago, Monden (1983) referred to cells (companies, etc.) being linked via *kanbans*, and Ohno (1988) referred to linking production lines (where each line may in fact be in the form of cells).

The card-based element of POLCA is described following the framework proposed in Liberopoulos & Dallery (2000). Since there is no output queue, there are only three elements. Queue A_i contains the POLCA cards for station i with $i = 1, \dots, n$ where n is the number of stations in the system. Queue P_0 is the queue of newly created jobs that are to enter the system and have no POLCA cards attached to them. This queue reflects the job arrival rate (or demand rate) λ . Finally, queue PA_i contains the jobs finished at the preceding station and to which a POLCA card from the preceding station is still attached. POLCA uses card-loops between pairs of stations. For example, a POLCA 1-2 card is used between Station 1 and Station 2. Imagine a shop that produces jobs that move from Station 1 to Station 2 to Station 3. This situation is illustrated in Figure 1. When a customer places an order, a new job is created and enters queue P_0 . The job waits in queue P_0 until a POLCA 1-2 card is available in queue A_1 . Once this card is available, the job is processed and moves to the queue PA_2 of the next station with the POLCA 1-2 card attached to it. The job waits in queue PA_2 until a POLCA 2-3 card is available in queue A_2 . Once this card is available, the job is processed. After processing, the POLCA 1-2 card is freed and moves back to queue A_1 and the job moves to the queue of the next station PA_3 with the POLCA 2-3 card attached to it. Thus, card loops are overlapping since the POLCA 1-2 card is only released after the operation at Station 2 has been completed.

[Take in Figure 1]

2.2 Mechanisms Underlying a POLCA System – Authorization

A major difference between POLCA and, for example, *kanban* systems is that POLCA cards are job anonymous (Riezebos, 2010; Ziengs *et al.*, 2012). In other words, POLCA cards do not indicate which job to work on – just that a job that requires a certain loop (e.g. 1-2) in its routing can be worked on. This introduces the need for another means of choosing between multiple alternative jobs waiting in the queue of a station. This means is provided by the release Authorization element

– the “A” in POLCA. POLCA’s authorization element uses earliest release dates for each station calculated by a HL/MRP system. However, different to a ‘classical’ priority dispatching rule (which would prioritize orders according to earliest release dates), POLCA’s authorization element prohibits a station from working on a job even if a card is available (and even if the station is starving) if the earliest release date has not yet been reached. In other words, when the order arrives at Station 1, three conditions have to be met to start the order:

- (i) Station 1 must be available;
- (ii) A POLCA 1-2 card (which circulates between the station pair 1 and 2) must be available, indicating the future availability (of capacity) at Station 2; and,
- (iii) The order must be authorized, i.e. the earliest release date for this order at Station 1 must have been reached (if there are multiple authorized orders waiting, the most urgent order is normally processed next).

While the POLCA literature consistently emphasizes the importance of using the Authorization element (e.g. Suri, 1998; Krishnamurthy & Suri, 2009; Riezebos *et al.*, 2009; Riezebos, 2010; Suri, 2010), we highlight its possible drawbacks. For example, if earliest release dates are too long, there may be needless starvation; and if earliest release dates are too short, prioritization may be jeopardized.

2.3 Discussion of the POLCA Literature

POLCA has remained largely unchanged since its introduction (Riezebos, 2010). One of the only improvements reported has been the introduction of color-coded cards by Pieffers & Riezebos (2006, cited in Riezebos, 2010) – stations are given a specific color, meaning each POLCA card consists of two colors. Meanwhile, Vandaele *et al.* (2008) presented an approach for setting the number of POLCA cards in accordance with expected demand in the context of an electronic POLCA system. More recently, Thüerer *et al.* (2017) assessed the impact of different combinations of card-allocation & dispatching rules and introduced a starvation avoidance mechanism. However, none of these studies critically assessed the impact of the authorization element.

Similarly, while there have been several case study implementations of POLCA reported in the literature (e.g. Krishnamurthy & Suri, 2009; Riezebos, 2010; Ajay Guru Dev *et al.*, 2017), the accuracy of the earliest release dates calculated by the HL/MRP system remains unclear; it is not even clear whether this element has been implemented in prior work. In other words, it remains unclear whether authors have actually implemented POLCA in its entirety. Meanwhile, simulation studies considering POLCA, such as by Lödding *et al.* (2003), Fernandes & Carmo-Silva (2006),

Germes & Riezebos (2010), Harrod & Kanet (2013), Farnoush & Wiktorsson (2013), Braglia *et al.* (2014), and Thürer *et al.* (2017), have neglected POLCA's authorization element and only modeled the card-based component. For example, Lödding *et al.* (2003) stated that no order was hindered at release if the current operation's authorization time had not yet passed. Meanwhile, Vandaele *et al.* (2008) applied the authorization element at the first station in the routing of orders only. This station acts as a gateway for an order to the shop and means every order released to the shop floor is authorized (Vandaele *et al.*, 2008: p. 193).

It therefore follows that the actual performance of POLCA has not been fully assessed. To the best of our knowledge, there exists no prior study assessing the performance of a (complete) POLCA system. Rather, authors use paired-cell overlapping loops of cards (which are very similar to a *kanban* system; Thürer, *et al.*, 2016) in combination with a dispatching rule. In response, our study began with the following research question:

RQ1: What is the impact of the release authorization element (that puts the "A" into "POLCA") on POLCA performance?

An exploratory study based on controlled simulation experiments is used to compare the performance of a complete POLCA system with the performance of POLCA without Authorization (i.e. POLC) in combination with a dispatching rule. The simulation model and experimental factors used in the study will be outlined next in Section 3.

3. Simulation Model

The primary objective of the simulation experiments is to assess the impact of POLCA's authorization element. Stylized standard models will be used to avoid interactions that may interfere with our understanding of the experimental factors. This approach is in line with previous simulation studies on POLCA (e.g. Germes & Riezebos, 2010, Ziengs *et al.*, 2012, Braglia *et al.*, 2014). While any individual shop in practice will differ in many aspects from our stylized environment, the models used in this study capture the job and shop characteristics of high variety make-to-order shops for which POLCA was designed (Riezebos, 2010). In other words: shops with high routing variability, processing time variability, and arrival time variability. The shop and job characteristics modeled in the simulations are first outlined in Section 3.1. How POLCA with and without Authorization (POLCA vs. POLC) is modeled is then outlined in Section 3.2. Finally, the experimental design is outlined and the measures used to evaluate performance are presented in Section 3.3.

3.1 Overview of Modeled Shop and Job Characteristics

Two shops have been implemented using ARENA simulation software. Each shop consists of three production levels with 1, 2 and 4 stations, respectively. Each station is a single, constant capacity resource. While Germs & Riezebos (2010) and Ziengs *et al.* (2012) only considered divergent production flows, we consider divergent *and* convergent production flows since the release list is a prerequisite for the use of POLCA in convergent structures (Riezebos, 2010). The divergent structure refers to a type of plant that is known as the V-Plant, while the convergent structure refers to the so-called A-Plant (Umble, 1992). The two shop structures are illustrated in Figure 2a and 2b, respectively. Each order in the experiments has one operation at each production level, so each job has three operations. There are four possible routings (i.e. the sequence in which stations have to be visited by a job); and arrows are used to indicate the transition possibilities of jobs. All possible stations have an equal probability of being the next station in the routing of an order. There are no feedback loops in the routing of orders, which avoids the issue of blocking that has previously been observed for POLCA (Lödding *et al.*, 2003; Harrod & Kanet, 2013).

[Take in Figure 2]

Operation processing times follow a lognormal distribution, as advocated by Trietsch *et al.* (2012). Processing time variability is a factor that is likely to influence performance in practice. Therefore, three levels of processing time variability are modeled, with a squared coefficient of variation, $cv^2 = 0.25, 0.5, \text{ and } 1$. The level of 1 is equal to the variability of an exponential distribution, which is typically considered to represent “high” processing time variability. The “medium” level of 0.5 is equal to the variability of a 2-Erlang distribution. Finally, the level of 0.25 has been chosen to represent “low” processing time variability. This level is still sufficient to avoid unrealistic, nearly symmetric distributions, as observed for lower cv^2 levels. Set-up times are considered part of the operation processing time. The same distribution is used at all production levels, but to ensure equal average utilization levels at each production level, processing times are multiplied by 2 for stations 2 and 3 (the production level with 2 stations) and by 4 for stations 4, 5, 6, and 7 (the production level with 4 stations). Similar levels of average utilization across stations are required to avoid stationary bottlenecks, which would distract attention from the main focus of our study.

The inter-arrival time of jobs follows an exponential distribution with a mean of 1.111, which based on our routing and processing time characteristics deliberately results in a utilization level of 90%. Due dates are set exogenously by adding a random allowance factor, uniformly distributed

between 75 and 100 time units, to the job entry time. The minimum value will be sufficient to avoid that a job's total processing time exceeds the allowance. The maximum value was chosen through preliminary simulation experiments such that it would result in a percentage of tardy jobs that is neither too low (so our results are not affected by incidental effects, as very few jobs would be responsible for the performance of the shop) nor too high (so certain adverse effects are avoided, since control methods that reduce the variance of lateness across jobs might even lead to an increase in the percentage tardy when due date allowances are too tight on average) for both investigated POLCA systems (POLCA and POLC, i.e. POLCA with and without authorization).

3.2 Production Control

As in Germs & Riezebos (2010) and Ziengs *et al.* (2012), POLCA card loops are established for each routing step. Two POLCA systems are considered: POLCA (see Section 2 above) and POLC. POLC is a version of POLCA that disregards POLCA's authorization element, i.e. it only uses POLCA's paired-cell overlapping loops of cards (which are then supported by a dispatching rule instead of the authorization element). The difference between POLCA and POLC is that, for POLCA, a job can *only* start at a station if its earliest release date has passed whereas, for POLC, the earliest release date is used for priority dispatching. Priority dispatching means that the earliest release date determines the sequence in which jobs are processed, irrespective of whether this earliest release date has passed or not. Hence, the earliest release date does not restrict the set of eligible jobs.

For both systems – POLCA and POLC – earliest release dates are determined by backward scheduling operation throughput time allowances for each operation in the routing of a job from the due date. Vandaele *et al.* (2008), in the context of an electronic POLCA system, used a queuing model and a so-called advanced resources planning (ARP) system to dynamically calculate these allowances. In this study, we focus on POLCA as a card-based system, i.e. as it was originally designed. As suggested for POLCA (Riezebos, 2010), and as is typical in the literature on modeling MRP systems (e.g. Krishnamurthy *et al.* 2004, Steele *et al.* 2005, Jodlbauer & Huber, 2008), we use a planned operation throughput time that is offset at each level. This allowance is given by the cumulative moving average realized in a simulation experiment. In practice, there will typically also be an error in the estimation of operation throughput times. Therefore, to better understand the sensitivity of POLCA to inaccuracies in the estimated earliest release dates, five levels of estimation accuracy of the allowance are considered. This is modeled by decreasing/increasing the allowance by -20%, -10%, 0%, 10%, and 20%.

Finally, seven levels for the number of cards are considered: 10, 12, 14, 16, 18, 20 cards per loop, and infinite cards. As in Germs & Riezebos (2010), the same number of cards is used within each control loop. While this is a simplification, it is justified by the fact that both POLCA and POLC share the same card loops. Thus, we do not seek to optimize the system for a certain performance measure but rather to focus on the performance differences. For an assessment of the impact of using a variable number of cards, the reader is instead referred to Ziengs *et al.* (2012).

3.3 Experimental Design and Performance Measures

The experimental factors are summarized in Table 1. A full factorial design was used with 420 scenarios, where each scenario was replicated 100 times. All results were collected over 13,000 time units following a warm-up period of 3,000 time units. These parameters allow us to obtain stable results while keeping the simulation run time to a reasonable level.

[Take in Table 1]

Four main performance measures are considered in this study as follows: mean *total throughput time* – the mean of the completion date minus the job arrival date across jobs; *percentage tardy* – the percentage of jobs completed after the due date; *mean tardiness* – that is $T_j = \max(0, L_j)$, with L_j being the lateness of job j (i.e. the actual delivery date minus the due date of job j); and, the *standard deviation of lateness*.

The total throughput time is used as the main indicator of the balancing capabilities of the approaches being tested. It also reflects the average lateness of jobs, which can be derived directly from this measure (it is equal to the realized mean total throughput time minus the mean of the delivery time allowance). The main performance indicators in a make-to-order shop, where the customer waits for his/her order, are delivery speed and delivery reliability. Since we assume due dates are given by the customer, the major performance indicator is therefore delivery reliability. The main indicator of delivery performance in our study is consequently the percentage of tardy jobs, which is influenced by both the average lateness and the dispersion of lateness across jobs. To a degree, both the mean tardiness and the standard deviation of lateness measure the dispersion of lateness across jobs. The standard deviation has the advantage of being relatively independent of the mean lateness, while the mean tardiness can be strongly correlated with the mean lateness. The standard deviation of lateness is however more sensitive to extreme values than the mean tardiness. The standard deviation of lateness can also be used to indicate earliness. But although reduced earliness is the main justification for the use of the authorization element in the POLCA

literature (e.g. Vandaele *et al.*, 2008; Suri, 2010), reduced earliness is not considered a main performance measure in our study. This is because ‘perfect’ earliness can be achieved by delaying all jobs, which is not advisable; earliness has to be interpreted in the context of delivery reliability.

Finally, in addition to the four main performance measures, we also measure the average shop floor throughput time as an instrumental performance variable. While the total throughput time includes the time that an order waits before being released, the shop floor throughput time only measures the time after an order has been released to the shop floor. The average shop floor throughput time is a useful indicator of the work-in-process level on the shop floor as, according to Little’s law (Little, 1961), it is directly linked to the level of work-in-process.

4. Results

Statistical analysis has been conducted to obtain a first indication of the relative impact of the experimental factors. An Analysis of Variance (ANOVA) has been used since it can be considered reasonably robust for different input distributions (Schmider *et al.*, 2010). The ANOVA is here based on a block design, which is typically used to account for known sources of variation in an experiment. In our ANOVA, we treat both shop structure and processing time variability as blocking factors. This allows the main effects of both factors and the main and interaction effects of our other three POLCA related factors (POLCA system, allowance accuracy, and number of POLCA cards) to be captured. The results are presented in Table 2.

[Take in Table 2]

Most main effects and two-way interactions were shown to be statistically significant. There are also significant three-way interactions in terms of the percentage tardy. Meanwhile, the main effect of allowance accuracy and the two-way interaction between POLCA system and allowance accuracy are equivalent for all performance measures since POLC is not affected by different levels of the allowance accuracy – the priority for jobs is equal for the different levels of the accuracy factor (-20%, -10%, 0%, 10% and 20%). This also explains why the two-way interaction between allowance accuracy and the number of POLCA cards is equivalent to the three-way interaction.

In addition to the above, the Scheffé multiple comparison procedure was applied to obtain a first indication of the direction and size of the performance differences between POLCA and POLC. The results in Table 3 suggest that POLC has the potential to outperform POLCA in terms of the total throughput time, percentage tardy and mean tardiness while POLCA outperforms

POLC in terms of the standard deviation of lateness. To further assess performance differences, detailed performance results will be presented next in Section 4.1 where we focus on a divergent shop and medium processing time variability. The impact of our two blocking factors, shop structure and processing time variability, is then assessed in Section 4.2.

[Take in Table 3]

4.1 Performance Assessment: POLCA vs. POLC in Divergent Shops

To aid interpretation, the results are presented in the form of performance curves. The left-hand starting point of the curves represents the lowest number of cards allowed in a POLCA or POLC loop (10 cards). The number of cards allowed increases step-wise by moving from left to right in each graph, with each data point representing one card level (10, 12, 14, 16, 18, and 20 cards and infinite). Increasing the number of cards increases the level of work-in-process and, as a result, increases the shop floor throughput times. Figures 3a, 3b, 3c, and 3d show the total throughput time, percentage tardy, mean tardiness, and standard deviation of lateness results over the shop floor throughput time results, respectively in the divergent shop. Only results for the divergent shop and medium processing time variability are shown, with the impact of shop structure and processing time variability assessed in the next section.

[Take in Figure 3]

The following can be observed from the results:

- *POLCA vs. POLC*: POLC significantly outperforms POLCA in terms of the total throughput time, percentage of tardy jobs, and mean tardiness. Improved performance is realized at a much lower shop floor throughput time and thus level of work-in-process. POLCA only leads to better performance in terms of the standard deviation of lateness (since it reduces the earliness of jobs). This means that, contrary to the arguments put forward in the POLCA literature, POLCA's authorization element has a direct detrimental effect on performance. Whether this effect is really due to the authorization element can be verified by observing the result to the right of each curve. These points represent the results with an infinite number of cards, i.e. where only POLCA's authorization element is active.
- *Allowance Accuracy*: Only one level of allowance accuracy is given for POLC since the different levels resulted in equal dispatching priorities in our modeled environment. For POLCA, an underestimation of the allowance leads to a reduction in shop floor throughput times (i.e. a shift to the left, given that this measure is on the x-axis), while an overestimation

leads to an increase (i.e. a shift to the right). Meanwhile, overestimating allowances leads to faster releases to the shop floor (since the release date for the first station is earlier). The resulting reduction in pool waiting times (total time minus shop floor throughput time) outweighs the increase in shop floor throughput times, which leads to shorter total throughput times. This in turn leads to improved percentage tardy and mean tardiness performance. However, the effect of the allowance accuracy is much weaker than the direct detrimental effect of POLCA's authorization element.

4.2 Robustness Analysis: The Impact of Shop Structure and Processing Time Variability

Similar conclusions on the relative performance of POLCA vs. POLC to those in the divergent shop can be drawn from our results in the convergent shop. Thus, the findings can be considered to be robust to the shop structure. This is illustrated by Figures 4a, 4b, 4c, and 4d, which show the total throughput time, percentage tardy, mean tardiness, and standard deviation of lateness results over the throughput time results, respectively in the convergent shop.

[Take in Figure 4]

The main difference that can be observed by comparing Figure 3, from the divergent shop, with Figure 4, from the convergent shop, is a weaker detrimental performance effect if the number of POLCA cards is reduced (i.e. by moving from right to left in the figures). Meanwhile, our findings are also robust to processing time variability. This can be observed from Figure 5 and Figure 6, which give the results obtained for low and high processing time variability in the divergent and convergent shop, respectively. The main difference that can be observed by comparing Figure 4a (low) with Figure 4b (high) and Figure 5a (low) with Figure 5b (high) is that performance difference between POLCA and POLC diminish. This is due to general performance deterioration; results suggest that in shops with high processing time variability neither POLCA nor POLC should be applied.

[Take in Figure 5 & Figure 6]

5. Conclusions

POLCA is an important production planning and control concept in the literature that was developed to support the adoption of a Quick Response Manufacturing approach (e.g. Suri, 1998 and 2010). It has been argued to be an alternative to *kanban* systems specifically for high variety make-to-order production. POLCA is unique since it combines a card-based element (the paired

cell overlapping loops of cards) with a higher level Material Requirements Planning (HL/MRP) system for release authorization. In POLCA, an order is not authorized for production at a station if its earliest release date calculated by the HL/MRP system has not been reached. Yet although POLCA has been widely applied in research, the role of its integral authorization element has, to the best of our knowledge, never been assessed. Although the POLCA literature has emphasized the importance of this element and outlined the positive effect that it would have on performance, it appears as though POLCA is ultimately operationalized without this element. Prior case studies and simulations make no explicit reference to calculating and using earliest release dates from an HL/MRP system. In response, this study started by asking: What is the impact of the release authorization element (which puts the “A” into “POLCA”) on POLCA performance? In contrast to the arguments put forward in the POLCA literature, our results strongly argue against the use of POLCA’s authorization element. This has important implications for managers and research.

5.1 Managerial implications

The main implication of this study is that POLCA is likely to be more successful without its authorization element. This is a significant refinement to the POLCA system typically described in the literature. Implementing the authorization element may reduce earliness, but this is likely to be achieved at the expense of a large increase in the percentage tardy and the mean tardiness. Our results suggest that POLCA’s card-based element (POLC) – which controls the use of capacity at stations – is better combined with a dispatching rule for prioritization locally at stations on the shop floor, as is typically the case for MRP systems (Vandaele *et al.*, 2008). In fact, we did not encounter any implementation in the literature that mentions the authorization element as part of an implementation. It is therefore likely that most performance gains obtained from the use of POLCA in practice are attributable to POLC. Meanwhile, confusion between the full POLCA system described in the literature and the narrower systems actually used in practice may have led to failed implementations and hindered the wider adoption of POLCA. We therefore hope that our refinement furthers the adoption of this important production control concept.

5.2 Limitations and Future Research

A major limitation of our study is that it is based on controlled simulation experiments. While this is a legitimate approach since we require controlled experiments to clearly single out performance effects, future research could seek to identify failed POLCA implementation attempts to further corroborate our findings. Similarly, future research could explore the major obstacles to the

adoption of QRM and its POLCA methodology. This would extend the recent study by Godinho Filho *et al.* (2017) that assessed the degree of QRM readiness and application of organizations.

References

- Ajay Guru Dev, C., Senthil Kumar, V.S., and Rajesh, G., 2017, Effective human utilization in an original equipment manufacturing (OEM) industry by the implementation of agile manufacturing: A POLCA approach, *Human Factors and Ergonomics in Manufacturing*, 27, 2, 1, 79-86.
- Braglia, M., Castellano, D., and Frosolini, M., 2014, Optimization of POLCA-controlled production systems with a simulation-driven genetic algorithm, *International Journal of Advanced Manufacturing Technology*, 70, 385 – 395.
- Farnoush, A., and Wiktorsson, M, 2013, POLCA and CONWIP performance in a divergent production line: an automotive case study, *Journal of Management Control*, 24, 159 – 186.
- Fernandes, N.O., Land M.J. and Carmo-Silva, S., 2016, Aligning Workload Control Theory and Practice: Lot Splitting and Operation Overlapping Issues, *International Journal of Production Research*, (in print).
- Fernandes, N.O. and Carmo-Silva, S., 2006. Generic POLCA - A production and materials flow control mechanism for quick response manufacturing, *International Journal of Production Economics*, 104, 1, 74-84.
- Germes, R., and Riezebos, J., 2010, Workload balancing capability of pull systems in MTO production, *International Journal of Production Research*, 48, 8, 2345-2360.
- Godinho Filho, M., Marchesini, A.G., Riezebos, J., Vandaele, N., Ganga, G.M.D., 2017, The application of Quick Response Manufacturing practices in Brazil, Europe, and the USA: An exploratory study, *International Journal of Production Economics*, 193, 437–448.
- Gonzalez, P.L., Framinan, J.M, Pierreval, H., 2012, Token-based pull production control systems: an introductory overview, *Journal of Intelligent Manufacturing*, 23, 5-22.
- Harrod, S., and Kanet, J.J., 2013, Applying work flow control in make-to-order shops, *International Journal of Production Economics*, 143, 620-626.
- Jodlbauer, H. and 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, 8, 2179–2195.
- Kanet, J.J., 1988, Load-limited order release in job shop scheduling systems, *Journal of Operations Management*, 7, 3, 44 – 58.

- Krishnamurthy, A., Suri, R., and Vernon, M., 2004. Re-Examining the Performance of MRP and Kanban Material Control Strategies for Multi-Product Flexible Manufacturing Systems, *International Journal of Flexible Manufacturing Systems*, 16, 1, 123–150.
- Krishnamurthy, A. and Suri, R., 2009, Planning and Implementing POLCA: a card-based control system for high variety or custom engineered products, *Production Planning & Control*, 20, 7, 596-610.
- Land, M.J., Stevenson, M., Thürer, M., and Gaalman, G.J.C., 2015; Job Shop Control: In Search of the Key to Delivery Improvements, *International Journal of Production Economics*, 168, 257-266.
- Land, M.J., and Gaalman, G.J.C., 1998, The performance of workload control concepts in job shops: Improving the release method, *International Journal of Production Economics*, 56-57, 347-364.
- Liberopoulos, G., and Dallery, Y., 2000, A unified framework for pull control mechanisms in multi-stage manufacturing systems, *Annals of Operations Research*, 93, 1-4, 325-355
- Little, J., 1961, A proof of the theorem $L = \lambda W$, *Operations Research* 8, 383-387.
- Lödding, H., Yu, K.-W., and Wiendahl, H.-P., 2003, Decentralized WIP-oriented manufacturing control (DEWIP), *Production Planning & Control*, 14, 1, 42-54.
- Monden, Y., 1983, *Toyota Production System: Practical Approach to Production Management*, Industrial Engineering and Management Press, Norcross, Georgia.
- Ohno, T., 1988, *Toyota Production System: Beyond Large-Scale Production*, 1st Ed., Productivity Press.
- Riezebos, J., 2010, Design of POLCA material control systems, *International Journal of Production Research*, 48, 5, 1455-1477.
- Riezebos, J., Klingenberg, W., and Hicks, C., 2009, Lean Production and information technology: Connection or contradiction?, *Computers in Industry*, 60, 237 – 247.
- Severino, M.R., and Godinho Filho, M., 2017, POLCA system for supply chain management: simulation in the automotive industry, *Journal of Intelligent Manufacturing*, DOI 10.1007/s10845-017-1323-5
- Schmider, E., Ziegler, M., Danay, E., Beyer, L., & Bühner, M., 2010, Is it really robust? Reinvestigating the Robustness of ANOVA Against Violations of the Normal Distribution Assumption, *Methodology*, 6, 4, 147-151.
- Stalk, G., 1990, *Competing against time: How time-based competition is reshaping global markets*. Simon and Schuster.

- Steele, D.C., Philipoom, P.R., Malhotra, M.K., and Fry T.D., 2005, Comparisons between drum-buffer-rope and material requirements planning: a case study, *International Journal of Production Research*, 43, 15, 3181-3208.
- Stump, B., and Badurdeen, 2012, Integrating lean and other strategies for mass customization manufacturing: a case study, *Journal of Intelligent Manufacturing*, 23, 109-124.
- Suri, R., 1998, *Quick Response Manufacturing: A companywide approach to reducing leadtimes*, Productivity Press.
- Suri, R., 2010, *It's about time: The competitive advantage of quick response manufacturing*, Productivity Press.
- Thürer, M., Fernandes, N.O., Carmo-Silva, S., and Stevenson, M., 2017, Improving Performance in POLCA Controlled High Variety Shops: An Assessment by Simulation, *Journal of Manufacturing Systems*, 44, 143-153.
- Thürer, M., Stevenson, M., and Protzman, C.W., 2016, *Card-based Control Systems for a Lean Work Design: The Fundamentals of Kanban, ConWIP, POLCA and COBACABANA*, Productivity Press, ISBN 978-1498746946.
- Thürer, M., Stevenson, M., Silva, C., Land, M.J., and Fredendall, L.D., 2012, Workload control (WLC) and order release: A lean solution for make-to-order companies, *Production & Operations Management*, 21, 5, 939 – 953.
- Trietsch, D., Mazmanyan, L., Gevorgyan, L., Baker, K.R., 2012, Modeling activity times by the Parkinson distribution with a lognormal core: Theory and validation, *European Journal of Operational Research* 216, 2, 386-396.
- Umble, M., 1992, Analyzing Manufacturing Problems Using V-A-T Analysis, *Production and Inventory Management Journal*, 2nd Quarter, 55 - 60.
- Vandaele, N., Van Nieuwenhuyse, I., Claerhout, D., and Cremmery, R., 2008, Load-Based POLCA: An Integrated Material Control System for Multiproduct, Multimachine Job Shops, *Manufacturing & Service Operations Management*, 10, 2, 181-197.
- Ziengs, N., Riezebos, J., & Germs, R., 2012, Placement of effective work-in-progress limits in route-specific unit-based pull systems, *International Journal of Production Research*, 50, 16, 4358-4371.

Table 1: Summary of Simulated Shop and Job Characteristics

Shop structure	Divergent and convergent flows
Processing time variability	low ($cv^2 = 0.25$), medium ($cv^2 = 0.5$) and high ($cv^2 = 1$)
POLCA system	POLCA and POLC
Allowance accuracy	deviation factor of -20%, -10%, 0%, 10% and 20%
Number of POLCA cards	10, 12, 14, 16, 18, 20 and infinite cards

Table 2: ANOVA Results

	Source of Variance	Sum of Squares	Degree of Freedom	Mean Squares	F-Ratio	p-Value
Total throughput time	Shop Structure (S)	3867541.50	1	3867541.50	4213.12	0.00
	Processing Time Variability (P)	23723734.00	2	11861867.00	12921.77	0.00
	POLCA System (POLC-A)	24422546.00	1	24422546.00	26604.79	0.00
	Accuracy (A)	278111.27	4	69527.82	75.74	0.00
	Number of Cards (N)	10457854.00	6	1742975.70	1898.72	0.00
	POLC-A x A	278111.27	4	69527.82	75.74	0.00
	POLC-A x N	190299.44	6	31716.57	34.55	0.00
	A x N	4279.74	24	178.32	0.19	1.00
	POLC-A x A X N	4279.74	24	178.32	0.19	1.00
	Residual	38487955.00	41927	917.98		
Percentage tardy	Shop Structure (S)	23.85	1	23.85	2133.25	0.00
	Processing Time Variability (P)	581.42	2	290.71	25999.49	0.00
	POLCA System (POLC-A)	2282.19	1	2282.19	200000.00	0.00
	Accuracy (A)	46.69	4	11.67	1044.03	0.00
	Number of Cards (N)	182.71	6	30.45	2723.36	0.00
	POLC-A x A	46.69	4	11.67	1044.03	0.00
	POLC-A x N	2.85	6	0.47	42.47	0.00
	A x N	1.67	24	0.07	6.20	0.00
	POLC-A x A X N	1.67	24	0.07	6.20	0.00
	Residual	468.80	41927	0.01		
Mean tardiness	Shop Structure (S)	2742316.30	1	2742316.30	3407.53	0.00
	Processing Time Variability (P)	13219533.00	2	6609766.40	8213.13	0.00
	POLCA System (POLC-A)	2897027.70	1	2897027.70	3599.77	0.00
	Accuracy (A)	118153.88	4	29538.47	36.70	0.00
	Number of Cards (N)	7659237.70	6	1276539.60	1586.20	0.00
	POLC-A x A	118153.88	4	29538.47	36.70	0.00
	POLC-A x N	343405.35	6	57234.23	71.12	0.00
	A x N	581.02	24	24.21	0.03	1.00
	POLC-A x A X N	581.02	24	24.21	0.03	1.00
	Residual	33742028.00	41927	804.78		
Standard deviation (SD) of lateness	Shop Structure (S)	523705.92	1	523705.92	2135.92	0.00
	Processing Time Variability (P)	7459473.80	2	3729736.90	15211.64	0.00
	POLCA System (POLC-A)	24895.57	1	24895.57	101.54	0.00
	Accuracy (A)	640.94	4	160.24	0.65	0.62
	Number of Cards (N)	1585784.00	6	264297.33	1077.93	0.00
	POLC-A x A	640.94	4	160.24	0.65	0.62
	POLC-A x N	4159.79	6	693.30	2.83	0.01
	A x N	360.20	24	15.01	0.06	1.00
	POLC-A x A X N	360.20	24	15.01	0.06	1.00
	Residual	10280067.00	41927	245.19		

Table 3: Results for Scheffé Multiple Comparison Procedure

Release Method (x)	Release Method (y)	Total Throughput Time		Percentage Tardy		Mean Tardiness		SD of Lateness	
		lower ¹⁾	upper	lower	upper	lower	upper	lower	upper
POLCA	POLC	-48.81	-47.65	-0.47	-0.46	-17.15	-16.07	1.24	1.84

¹⁾ 95% confidence interval; * not significant at $\alpha=0.05$

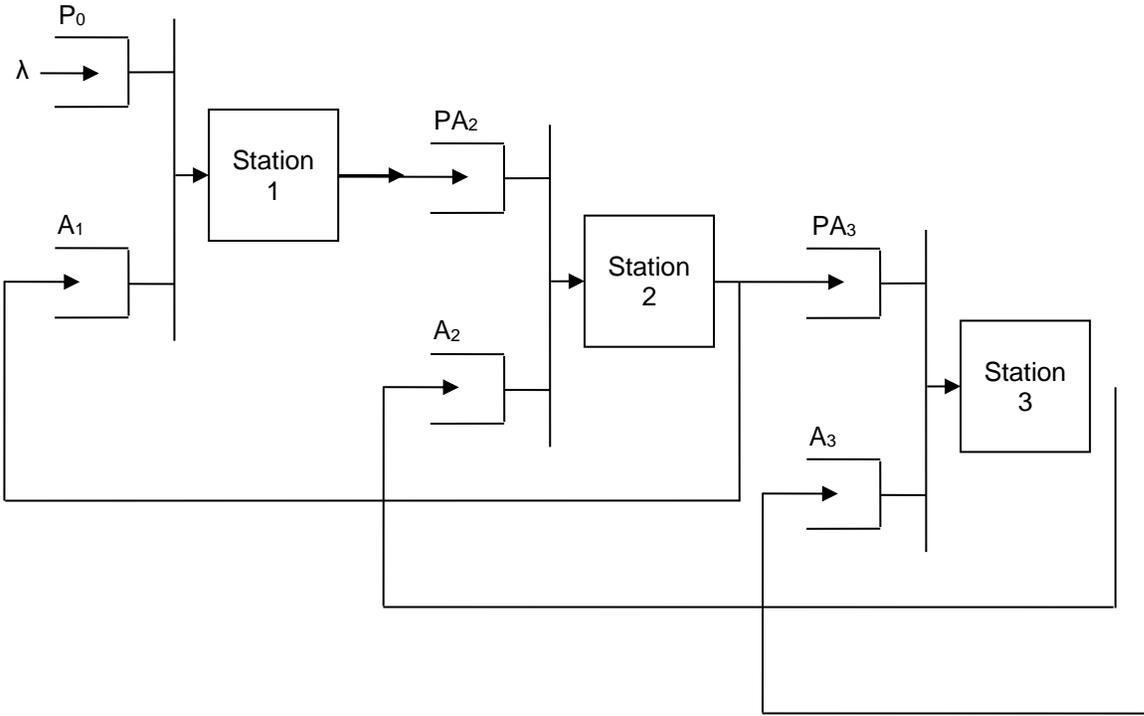


Figure 1: POLCA's Card-Based Control Loops

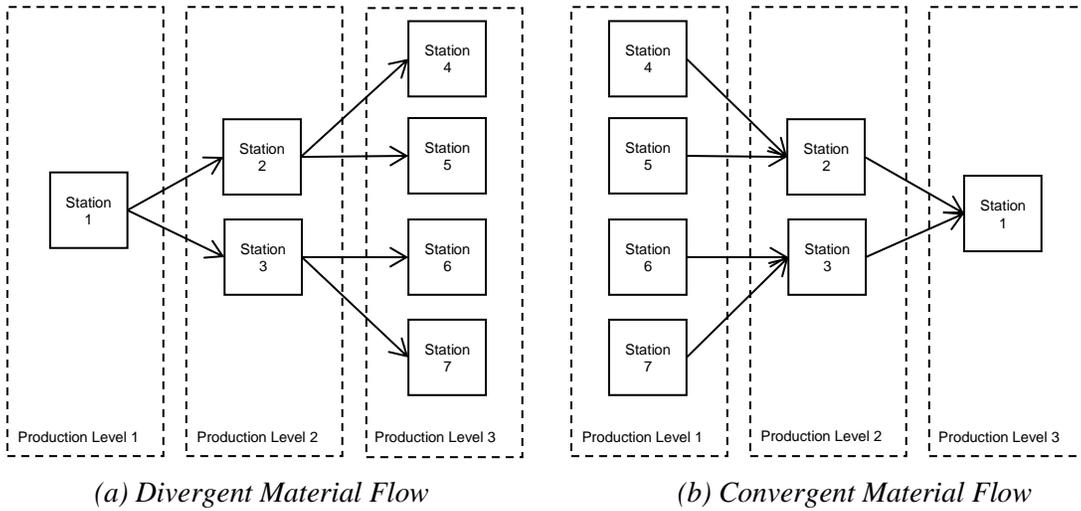


Figure 2: Modelled Shop Configurations – Divergent and Convergent Material Flows

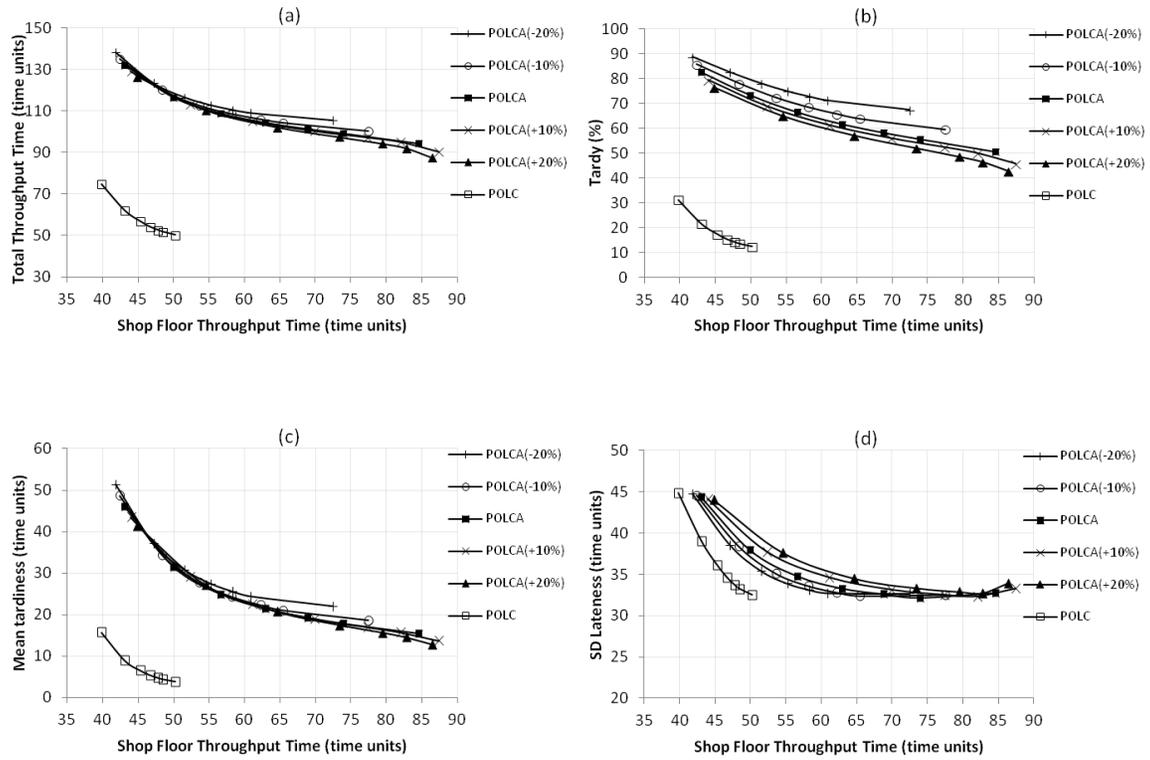


Figure 3: Performance Results for POLCA vs. POLC in the Divergent Shop (Medium Processing Time Variability) – (a) Total Throughput Time; (b) Percentage Tardy; (c) Mean Tardiness; and (d) Standard Deviation of Lateness over the Shop Floor Throughput Time

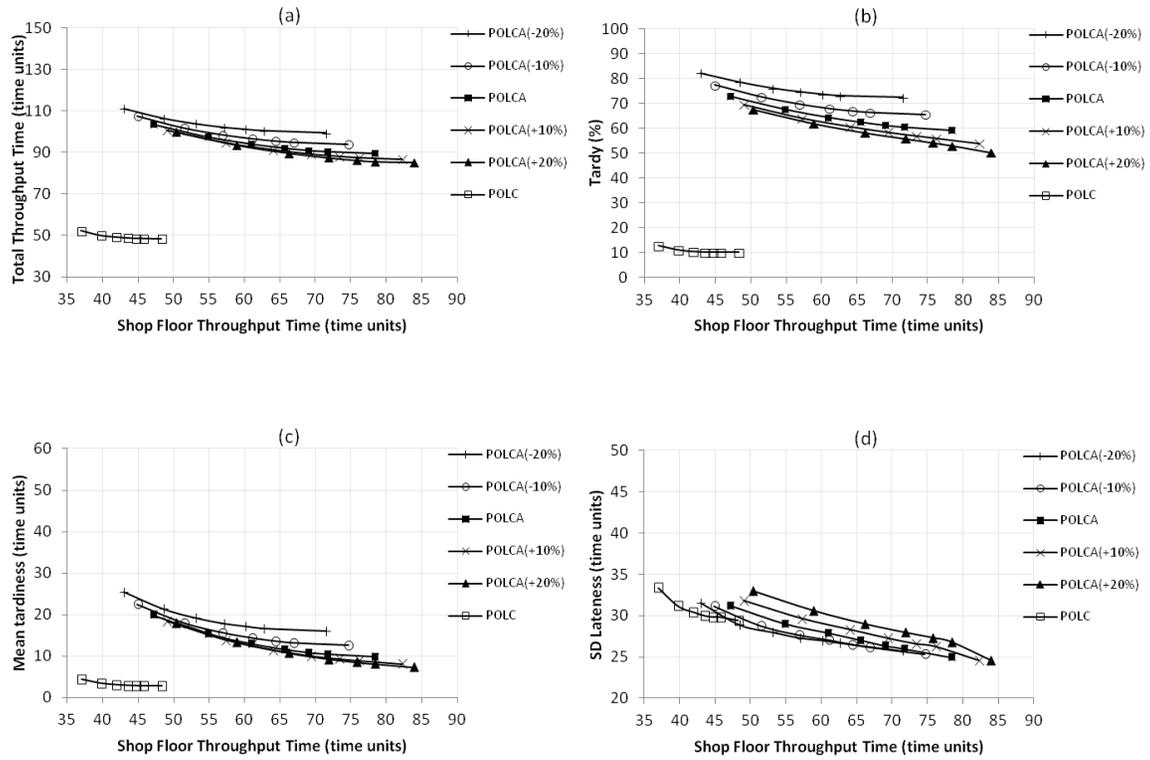
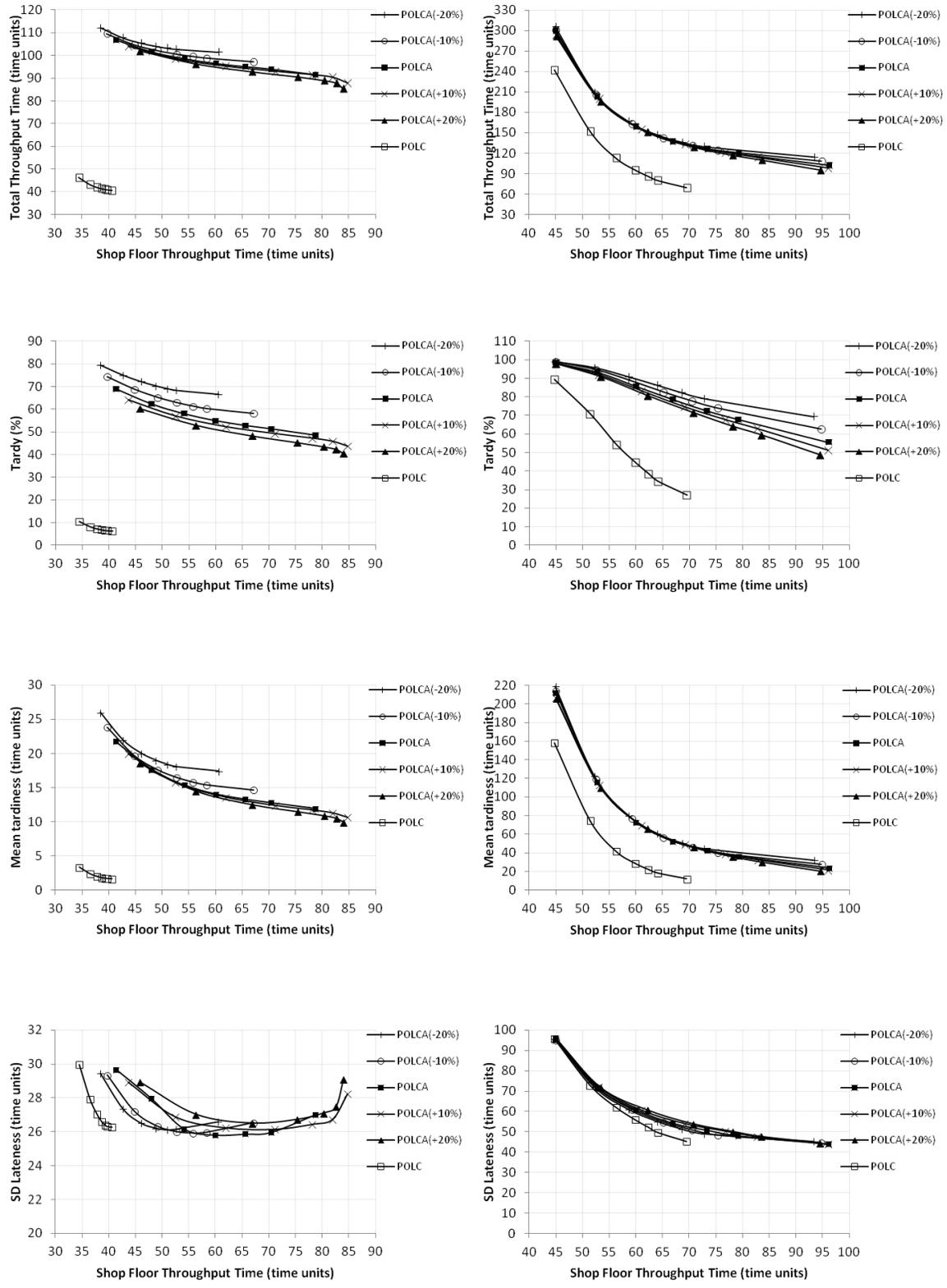


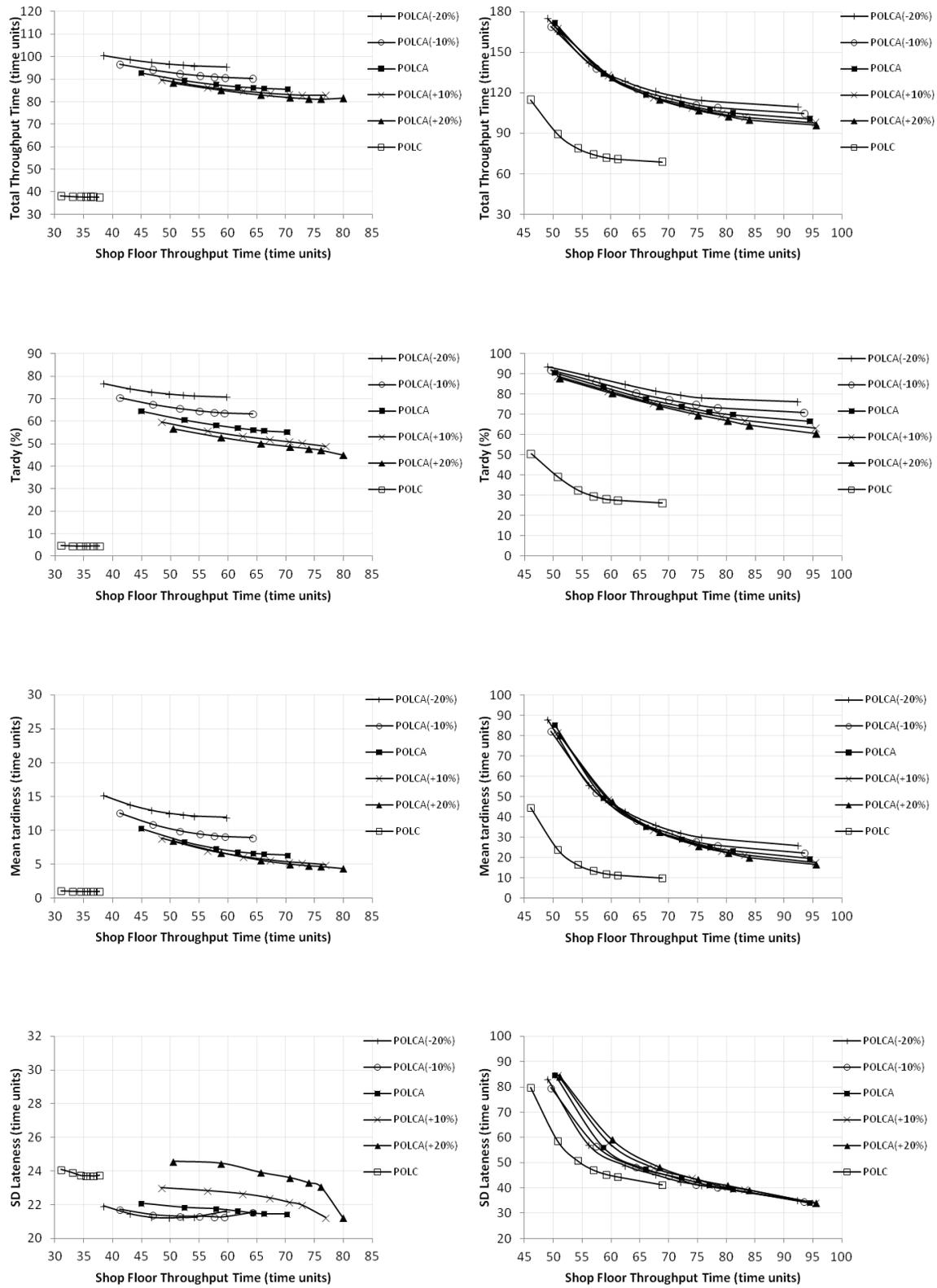
Figure 4: Performance Results for POLCA vs. POLC in the Convergent Shop (Medium Processing Time Variability) – (a) Total Throughput Time; (b) Percentage Tardy; (c) Mean Tardiness; and (d) Standard Deviation of Lateness over the Shop Floor Throughput Time



(a)

(b)

Figure 5: Performance Results for POLCA vs. POLC in the Divergent Shop – (a) Low Processing Time Variability; and (b) High Processing Time Variability



(a)

(b)

Figure 6: Performance Results for POLCA vs. POLC in the Convergent Shop – (a) Low Processing Time Variability; and (b) High Processing Time Variability