The Pennsylvania State University

The Graduate School

College of Engineering

WORKFORCE SCHEDULING IN THE CONTEXT OF HUMAN PERFORMANCE VARIABILITY: THE WORKER APPROACH

A Dissertation in

Industrial Engineering

By

Frank Bentefouet

Submitted in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

May 2013

The dissertation of Frank Bentefouet was reviewed and approved* by the following:

David A. Nembhard

Associate Professor of Industrial and Manufacturing Engineering

Dissertation Advisor, Chair of Committee

Christopher H. Griffin

Lead Research Associate, Cyber-Analytics and Operations Research Group

Applied Research Laboratory

Paul M. Griffin

Professor of Industrial and Manufacturing Engineering

Peter and Angela Dal Pezzo Department Head Chair

Soundar Kumara

Pearce Chair Professor of Industrial and Manufacturing Engineering

Professor of Computer Science and Engineering

*Signatures are on file in the Graduate School

ABSTRACT

The combination of short product life cycle, intense competition, and global markets has increased the pressure on manufacturing and service industries. The broad spectrum of market challenges has forced businesses to move toward greater agility and flexibility. In addition to this agility and flexibility requirements, innovation has become a key factor for market share increase. As a consequence, processes involved in the conception of the ultimate goods have to undergo constant changes. For an organization, it is then critical to learn and implement those changes faster than its competitor in order to economically survive (Kapp, 1999). It has become clear that improvements should focus not only on the efficiency and effectiveness of technical processes but also on the workers involved in those processes. At the factory level, the ability of an organization to adapt relies for a great share on workers ability to learn and master tasks in short periods of time.

Facilities run on schedules that are often built in accordance with a defined manufacturing strategy; reduce delivery times, increase productivity rates, or minimize number of late jobs. However, when the classical theory of scheduling is applied to task sequencing problems, human characteristics are often not the primary consideration and their impact is often assumed away.

The goal of this dissertation is to contribute in reducing the gap between the theory of scheduling and conditions in factories. In the presented research, human variability is considered when addressing the worker-to-task allocation problem in serial and parallel systems with the aim of maximizing throughput. In the context of the thesis, the worker performance variability is explained by Learning and Forgetting (L&F) effects. Learning effect is modeled to occur when repeatedly performing the same task and forgetting effect is assumed to happen with breaks, mainly when the worker is away from the task. Two primary ways of learning are considered: learning by doing and learning from other workers by transfer of knowledge. The methodology includes human performance modeling, math program formulations of the worker-to-task problem and simulation of work organization.

The study of math formulations provides structural and analytical properties of optimal schedule in parallel production scenario. The structural property of the optimal schedule is proved to be relevant with stochastic learning, tasks similarities, and unequal number of workers and tasks. The analysis of the serial system mathematical formulation leads to the characterization of the optimal switching time between workers and to an equivalent tractable math formulation. The analysis allows delineating work sharing policies and specialization policy. Simulation of work organization provides managers with insights on cross-training level, multi-functionality, preferable level of specialist to generalist ratio, and their impact on system performance. Results demonstrate that when setting cross-training level, selecting workers based on prediction of system output is outperformed by other simpler prior expertise based policies. When modeling learning by doing and learning by transfer of knowledge, heuristics are suggested at the hiring stage, the grouping stage and the assignment stage. The suggested policies have the advantage to be straightforward to implement in the work place and do not rely on solving math programs. The heuristics performance illustrates the importance to assess and consider the ability of workers to learn from others when building schedules.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

My journey toward earning the doctorate degree started before my admission to the Ph.D. program. It began at Ecole Centrale Paris in France when, working on my master I attended in 2006 a presentation about studying at the Pennsylvania State University by Dr. David Nembhard who three years later became my thesis advisor. I would like to thank all the people whose positive influences have contributed in the making of that journey. My deepest gratitude goes to my advisor Dr. David Nembhard who made the Ph.D. quest possible. A summer internship in his Penn State Lab in 2007 generated an interest for the research and two years later he made it happen. He has been a constant help and a steady support through the years. Our weekly research meeting which was always a discussion from one investigator to another has created an ideal atmosphere for work. Research often comes with frustrations and as a young investigator, the need of someone motivating you is critical. I have been fortunate to have Dr. David Nembhard as a thesis advisor and I am deeply grateful to him. In my first semester, I registered for his workforce engineering course and it was a delightful introduction to the research on manpower scheduling and the challenge of building planning that integrates human performance characteristics. My gratitude also goes to my committee members; Dr. Christopher Griffin who is an open encyclopedia on optimization, he brought his expertise with energy and passion, Dr. Soundar Kumara for sharing his knowledge, he showed me that there are no boundaries in science and I hope his passion for learning has been contagious, Dr. Paul Griffin for the constant encouragements and his constructive critics.

I also thank Dr. Jose Ventura for his help while working on my research. I express my gratitude to Dr. M. Jeya. Chandra, he has been a constant source of answers for various administrative questions, research, course work and concerns that I had. His support has been critical in the completion of the Ph.D.

I was fortunate to have through the years of the doctoral studies, the kind attention of Ms Olga Covasa and the Penn State Industrial and Engineering Department staff. I had great colleagues in the Center for Service Enterprise Engineering lab. They were friendly, supportive and eager to help. Special thanks to Gretchen, Camila, Amir, Ke, and Pedro. I wish them the best.

I thank my brother and sisters Leslie, Linda, Fanny and Romuald, my aunt and uncle Mustapha Thomas and Caroline Lapnet, my brother in law Patrick Temfack, my longtime friends Hermann, Kassim, Biko, Linda, Houkay, Maureen, Erica and Larissa. I am also grateful to Monifa, Seth, and my friends Qais and Marta, they helped me adapt to a new country.

A special acknowledgment to the Industrial and Manufacturing Engineering Department of the Pennsylvania State University and to his Department Head Dr. Paul Griffin for providing funding support through my doctoral studies.

I dedicate this thesis, to my late grandparents Doh Djibrilla Pierre, Tavo Joseph and Josephine Rose nee Mohou Mebouette, they have always wanted the descendants to be highly educated, to my grandmother Geupa Monique who I need to spend more time with, and to my parents Janvier and Sabsia Marie-Claire Bentefouet for their love and for raising their children with high standards. They are my role models.

CHAPTER 1

INTRODUCTION

1.1. Motivation

"Human capital" is a key success factor nowadays and to achieve their objectives, firms must consider the characteristics of workers involved in the daily operations. With the increase in global competition, firms are more than ever under constant pressure to improve delivery times while maintaining high quality service. Shorter product life cycles have led to an increase in exposure, forcing organizations to not only seek improvements in the efficiency of technical processes but also on the workers involved in those processes. Tasks are changing and workers must be able to learn and master them in shorter periods of time. It is thus important to consider the most appropriate staffing when facing these challenges where managers may potentially be aided by developing codified approaches for hiring, training and assigning workers. Because of these dynamic workplace conditions and the inherent complexities of systems of humans and automated systems, it is essential that these methods acknowledge worker differences.

Production planning is commonly built on the assumption that all workers are identical in their abilities and their performances are stationary. In the meantime, the human factors literature claims that fundamental characteristics distinguish human from machine and influence their performance. The research literature on worker-to-task assignment, typically assumes that workers operate at a steady productivity level or at a set of discrete productivity levels. However, workers learn and forget tasks based on how often they are assigned to. The necessity of incorporating worker learning and relearning behavior has been noted by Wisner and Pearson (1993). Their study illustrates that worker-relearning behavior has a significant impact on labor assignment policy, especially when a worker is regularly required to transfer between tasks.

The current literature on the worker-to-task allocation problem applies to steady state behavior, that is, once learning and other transients have subsided. Extending the understanding of workforce dynamics, in particular toward shorter-term transient operations is of considerable

interest. It is also envisioned to address managerial worker-scheduling decisions during the transient phase when learning has a significant impact on individual and system performance. The scarce literature on worker-to-task assignment based on learning and forgetting behavioral characteristics can be explained by the difficulty in collecting detailed performance data at the individual level. Moreover, a potential significant limitation of many learning/forgetting models in the literature is that they commonly are designed as measurement tools and may not lend themselves well to use within other analytical or optimization frameworks. In fact, when individual worker learning and forgetting is incorporated, this introduces nonlinearity in the objective function. Broad managerial questions such as whether or not to cross-train and how often workers should switch tasks remain unanswered.

The dissertation aims to bridge the gap between human factors engineering and production planning using Operation Research (O.R.) techniques. Historically, schedules have been constructed with the aim to reduce variability in production systems, more precisely in tasks completion times. Two approaches were widely used: a task approach which stipulates that the major source of variability is the inherent variability of the tasks, and a workload approach which claims that the environment where the tasks are being performed explains the variability. While useful in some settings, the two approaches were unable to explain some of the field observations. Their limitations led to the emergence of a new hypothesis: The worker approach. With the worker approach, the major source of variability in task completion times is due to the worker performing the task. The new approach, by acknowledging workers differences provides better approximation when modeling the differential in tasks completion times. Workers differences are materialized by the non-homogeneity of productivity rates curves, which produces the differential in completion times and leads to variability in the system. Learning and forgetting phenomena have been discussed to explain the heterogeneity in worker's productivity curves.

1.2 Current work

The dissertation studies the allocation of worker-to-task when maximizing throughput. It acknowledges worker performance variability and workers differences. The proposed work adopts the worker approach concept and assumes that the variability in worker performance is the result of learning and forgetting effects. This thesis analyzes the implication of learning and

forgetting in workforce scheduling. Two types of production systems are considered: parallel system also denominated independent jobs and independent tasks, and serial system also called independent jobs and dependent tasks. The aim is to augment the extant knowledge on optimal scheduling in parallel and serial production systems when worker performance variability is due to learning and forgetting. The study has the goal to establish results that are not model specific.

In order to reach these goals, the dissertation adopts the following methodology: the worker-to-task allocation problem is formulated as a math program. The study of the mathematical formulations establishes structural/analytical properties on optimal conditions in the considered production scenarios when variability in human performance is included. The aim is to characterize optimal policy structure and conditions that could simplify the complexity of the math problem. Indeed, a less complex formulation has the advantage to: increase the scale of scheduling problems that can be solved to optimality, to provide insight on the objective value, and to implement and estimate under various workforce management constraints, policies that could appeal to managers. The remainder of the dissertation is organized as follows:

Chapter 2 is devoted to the literature review on personnel scheduling, learning and forgetting phenomena, and knowledge transfer (learning by transfer). The following chapters focus on establishing structural properties that reduce the complexity of the math scheduling problem in parallel system and serial system.

Chapter 3 establishes the structure of an optimal scheduling policy in independent tasks and independent jobs system when considering a general productivity model that includes learning and forgetting. It is proved that when maximizing throughput without demand requirements, specialization is an optimal policy. The proven optimal policy structure allows solving a Mixed Integer Linear Program (MILP), rather than a more complex Mixed Integer Non Linear Program (MINLP). As the presented methodology is independent of the productivity model, it has general applications irrespective of the specific learning/forgetting models employed. The chapter includes extensions to the case of tasks with dependent learning, and stochastic learning.

Chapter 4 discusses the optimal flow line policy to implement in the context of worker performance variability. A closed form solution for the optimal switching time between workers

in work-sharing systems is established in the case of two and three tasks with two workers, and a model for serial system with more tasks is suggested. For a general number of workers and tasks, the maximum number of changes between assignments is established as well as an equivalent expression of the optimal throughput. Later, the result presented in the chapter, allows estimating the performance of heuristics developed in the next chapters.

Chapter 5 tackles the problem of selecting workers in parallel and serial systems with cross-training levels and workforce composition. Attention is given to the interaction between the implemented policies and various factors such as the ratio of generalist workers to specialist workers, the level of multi-functionality, and the workforce heterogeneity. Results demonstrate that, when implementing the cross-training, selecting workers based on prediction of system output is outperformed by other simpler prior expertise based policies.

Chapter 6 aims to provide insights to managers by suggesting heuristic approaches for selecting, grouping and assigning workers to tasks when learning by doing and learning by transfer of knowledge are considered.

Chapter 7 concludes the work by summarizing the established results and explores future research that could complement the research presented here.

CHAPTER 2

LITERATURE REVIEW

2.1. Scheduling

A definition of a scheduling problem as well as its principal components can be found in every book in the literature. Sarin et al. (2010) define a scheduling problem as the determination of starting times for the jobs waiting to be processed at a single machine or multiple machines (resources) for the objective of optimizing an appropriate performance measure of interest. Pinedo (1995) put forward another definition; scheduling concerns the allocation of limited resources to task over time, a decision-making process that has, as a goal, the optimization of one or more objectives.

We define a scheduling problem as the organization over time of the execution of a set of tasks, taking into account time constraints as well as capability and capacity constraint on resources required for these tasks. A schedule constitutes a solution to the scheduling problem. It describes the execution of tasks and the allocation of resources over time with the aim of meeting one or more objectives. A pure technical view of scheduling has been dominant for many years, contributing to the development of combinatorial optimization – the main branch of mathematics employed in scheduling theory and the cornerstone of an extensive research literature on scheduling since the book by Baker (1974). McKay and Wiers (1999) cited ten key implicit assumptions in mathematical formulations. Among these ten assumptions, we can enumerate the following:

- The relatively small set of facts about scheduling problem handled by the mathematical model or algorithm is sufficient to capture the main characteristics of the problem.
- The objectives or measurement metrics do not vary over the time horizon being sequenced.
- The feasibility constraints defining the problem do not vary over the time horizon.

- Time can be modeled as an abstract time series with t(i), and t(j) and that sequencing is not dependent upon a Monday effect, a shift effect, a holiday effect, or a time of year effect.

The ability for an organization to have the right staff on duty at the right time when attempting to satisfy customers requirements, the shift of businesses to a more service oriented, and the potential benefits provided by optimized staff schedules have led to an increasing attention by researchers on personnel scheduling and rostering. Personnel scheduling and rostering can be defined as the process of constructing optimized work timetables for staff. In our context, we employ either personnel planning, personnel scheduling or rostering.

2.1.1. Problem classification and models of personnel scheduling

There have been several general surveys papers in the area of personnel scheduling, included Tien et al. (1982) and Bechtold et al. (1991). The latter concentrates on general labor scheduling models. Aggarwal (1982) proposed an early survey of some application areas and methodologies. Here, we adopt the classification proposed by Ernst et al. (2004); they suggest 6 stages associated with the processes of constructing a roster, starting from the determination of staffing requirements and ending with the specification of the work to be performed, over some time period, by each individual in the workforce. Although the subdivision into stages suggests a step by step procedure, the development of a particular roster may require only some of the stages and in many practical implementation, several of the stages may be combined into one procedure. The taxonomy is not intended to be complete, but aims to provide a general framework. Moreover, each stage is subject to requirements that depend on applications.

2.1.1.1. Demand modeling

At this stage, the goal is to determine how many staff is needed at different times over some planning period. The demand modeling is the process of translating a predicted pattern of incidents into associated tasks and then using the task requirements to ascertain a demand for staff. Incidents occur during the planning period and could be taken to mean inquiries to a center or a defined sequence of tasks. We can distinguish:

A *task based demand*: The demand is obtained from lists of individual tasks to be performed. Typical applications of task based demand are crew scheduling in airlines, railways, buses and other transportation systems.

A *Flexible demand*: The likelihood of future incidents is less well known and must be modeled using forecasting techniques.

A *Shift based demand*: In this scenario, the demand is directly obtained from a specification of the number of staff that are required to be on duty during different shifts. A shift is defined as a day's work for a worker. This type of demand is often the basis for nurse scheduling.

The literature on scheduling based on demand modeling is vast. Aardal and Ari (1987) discussed the determination of demand for staff in a manufacturing environment given a profile for cost and benefits. Atlason et al. (2004) studied the use of a hybrid simulation and integer programming approach when minimizing staffing cost in a call center subject to maintaining a specified service level over a specified number of time periods. Borst (2001) proposed a queuing model for calculating staff requirements in call center where multiple queues exist. Each queue was characterized by the skills required and its arrival rate.

2.1.1.2. Days off scheduling

Days off scheduling often arises in rostering problems with flexible and shift based demand. In their paper, Baker and Magazine (1977) exploited several versions of a days-off scheduling problem by considering different scenarios on days-off patterns and weekend patterns. A minimum workforce size is derived and optimal schedules were obtained. A significant advance in mathematical modeling and optimization of work-rest schedules is due to Bechtold et al. (1984). The optimal number, placement, and duration of rest periods during a fixed length time horizon was determined by a linear decay of work rate, and a linear recovery rate. Bertchold et al (1984) report an increase in worker productivity of 13% in a real-world environment.

2.1.1.3. Shift scheduling

Shift scheduling involves selecting a set of the best shifts from a (large) pool of candidate shifts on a single day. It is similar to crew scheduling in transportation systems. When rostering to flexible demand, we need to consider the timing of work and meal breaks, when scheduling to task based demand shift scheduling is usually called crew scheduling or crew pairing optimization. Atlason et al. (2002) proposed an integer formulation to determine minimum

staffing levels that are fed into a rostering problem. The queuing theory and simulation are combined to model the demand.

2.1.1.4. Line of work construction

The process of constructing a line of work depends on building blocks like shifts, duties or stints. We consider rules relating to lines of work that ensure the feasibility of individual lines of work, and the pattern of demand that ensures that, taken together with the rules relating to lines of work, we satisfy the work requirements at all times in the planning process. Line of work construction is often called tour scheduling when dealing with flexible demand, and crew rostering when dealing with crew pairings. Arthur and Ravindran (1981) describe a goal programming method for nurse rostering. Integer programming is used to solve the goal programming problem of determining days of work for the nurse. The objectives are formulated in terms of minimum staff levels, desired staff levels, nurse preferences and nurse special request. Hao et al. (2004) developed a neural network model for solving cyclic rostering problem for an airport ground staff. The problem was formulated as an integer linear program and converted into a neural network model. Numerical comparisons are made against simulated annealing, tabu search and genetic algorithms.

2.1.1.5. Staff assignment

Also called rostering assignment, it involves the assignment of individual staff to the lines of work. A line of work defines the complete work-schedule for an individual over the rostering horizon. Staff assignment is often done during construction of the work of lines. Goodale and Thompson (2004) study the case of individual assignment of workers with both heterogeneous productivity and cost of labor. The assignment problem is formulated as a set covering model but is solved using a simulated annealing method. Alvarez-valdes et al. (1999) investigated the problem of creating staff for aircraft refueling roster at airports in Spain. The problem is solved via a heuristic that includes the generation of all feasible weekly patterns and rest shifts then worklines are assigned to staff and finally actual shifts are assigned to each day.

2.1.1.6. Task assignment

Task assignments are often required when working shifts have been determined but tasks have not been allocated to individuals. It is the process of allocating a set of tasks, with specified start and end times and skill requirements, between a group of workers who have typically already

been assigned to a set of working shifts. Campbell (2002) described methods for solving the assignment arising from the need to assign cross-trained workers between a number of departments at the start of a shift. Tharmmaphornphilas and Norman (2004) discussed the problem of rotating workers between jobs so as to increase work variety, minimize the performance of repetitive tasks and exposure to noise. For different jobs performed over different time intervals, a job severity indices (JSI) is calculated and an integer program is used to determine rotation patterns that minimize the maximum JSI.

The earliest discussion on staff scheduling and rostering could be traced back to Edie's work on traffic delays at toll booths (1954), since staff scheduling and rostering method have been applied to various environments such as airlines, railways, health care systems, emergency services and manufacturing industries. Acknowledging the value of workforce and carefully considering the design and operation of a firm's workforce can substantially contribute to the improvement of the objectives of the firm. Hence, the human capital of a firm can constitute a key for its success. Workforce scheduling models typically do not describe the relationship between human characteristics and conditions, and their corresponding effects on performance and well-being. Our ensuing discussion will focus on workforce scheduling research with explicit human considerations.

2.1.2. Workforce scheduling

A workforce scheduling problem is generally concerned with determining an allocation of employees to work shifts that minimizes labor cost and satisfies constraints related to staffing/ service levels and labor laws/agreements. With the global competition, firms are under constant pressure to increase the level of customer satisfaction by improving quality, delivery speed and industrial processes. It has become clear that improvement on efficiency and effectiveness of technical process come along with a better management of workers involved in these processes. Since the early 1960's, the gap between theory and practice in scheduling has been noted and has been discussed by a number of researchers (Buxey, 1989; Dudek et al., 1992; Graves, 1981; MacCarthy and Liu, 1993; Pinedo, 1995; Rodammer and White, 1988; Crawford and Wiers, 2001). The lack of fit between the human element and the formal or systemized element found in the scheduling methods and systems has been speculated to be one of the reasons (McKay et al., 1989; Mckay and Wiers, 1999). The paper from McKay et al. (1988), on what was relevant in

the Job-Shop scheduling theory, has renewed the interest on the role and contribution of the human element in production scheduling. However, it is still under-researched compared to the effort placed on the mathematical and software aspects.

The early literature on the human factors in scheduling focuses on the persons who do the production scheduling, reviewing the role of these persons. In his seminal work, Dutton (1962) attempted to capture scheduling practice in a box manufacturer from a simulation model of scheduler behavior. Hurst and McNamara (1967) studied the decision rules of a planner in a textile mill, capturing them in an equation that could be used to decide machine combinations for unscheduled orders. Interest in the subject was renewed later in the 1980s, with a number of studies that have attempted to develop knowledge-based or expert systems to support planning processes (Fox and Smith, 1984; Kanet and Adelsberger, 1987). The work of McKay is noteworthy in the field (McKay et al, 1988 and McKay et al., 1995); interesting aspects include: the importance of scheduler's information network, the impact of performance criteria on scheduler behavior, the importance of scheduler expertise and experience is noted.

2.1.2.1. Operations research approaches to workforce scheduling

The literature on papers using operation research techniques to model and solve workforce scheduling problems with objective functions and constraints that emphasize ergonomic criteria is scarce. Nembhard (2001) developed a heuristic approach for assigning workers to task based on individual learning rates. The developed heuristic significantly improved overall productivity under empirically observed conditions and under many experimental conditions. Kostreva et al. (2002) model human factors in generating shift schedules. In their work, a simulation environment is constructed to evaluate the effectiveness of several shift scheduling strategies. The natural circadian rhythm as well as circadian rhythms induced by various work schedules, is simulated with a system of partial differential equations. The measure of the effectiveness of a given scheduling strategy is quantified by the deviation of its induced circadian rhythm from naturally occurring (simulated) circadian rhythm. The simulation environment was used to find the best scheduling strategy. Corominas et al. (2010) have studied the task assignment problem when work performance depends on experience of all tasks involved. The nonlinearity of their performance function and thus of the scheduling problem, was overcome by using a linearization technique. In their work, the main objective is to minimize the completion times. Other

researchers emphasize the impact of workforce schedules on ergonomic criteria, but without human characteristics. Chen and Yeung (1992) used goal programming and an expert system in an integrated approach to solve a nurse scheduling problem. Kostreva and Jennings (1991) and Malladi and Jo Min (2004) introduced an algorithm for generating shift schedules that accounts for circadian rhythms as well as other constraints as time off between shifts, the number of consecutive workdays, and the number of distinct shifts during any two-week period. Malladi and Jo Min (2004) used the Analytic Hierachy Process (AHP) (Saaty, 1980) to assign weights to various scheduling schemes. The ergonomic criteria are taken into account in terms of forward vs. backward shift rotation. A larger penalty was assigned by the weighting AHP schemes to schedules that were more disruptive to circadian rhythms and therefore had a higher ergonomic cost. Weights from the AHP method were incorporated into an integer programming model whose solution is used to construct an optimal shift scheduling policy.

2.1.2.2 Human characteristics and their influence on human performance

When classical scheduling theory is applied to human task sequencing problems, human characteristics are not the primary consideration and their impact on these problems is often assumed away. In the meantime, the engineering research literature on human factors indicates that humans have unique characteristics that separate them from their inanimate counterparts (Wickens et al., 2004, Chapter 16). Therefore, assumptions related to resources in traditional scheduling are not applicable when tasks are performed by human resources. Table 2.1 summarizes several assumptions that are usually made to simplify human behavior (Bourdreau et al., 2003).

In the context of this work, we contend that human-related characteristics play an important role in addressing the human task sequencing problem. Several factors, occurring individually or in combination, can have an adverse effect on the productivity of human completing a sequence of tasks.

Table 2. 1. Assumptions used to simplify human behavior in OM models*

1. People are not a major factor (many models look at machines without people, so the human side is omitted entirely).
2. People are deterministic and predictable. People have perfect availability (no breaks, absenteeism, etc.). Task times are deterministic. Mistakes do not happen, or mistakes occur randomly. Workers are identical (work at the same speed, have the same values, respond to the same incentives).
3. Workers are independent (not affected by each other, physically or psychologically).
4. Workers are "stationary". No learning, tiredness, or other patterns exist. Problem solving is not considered.
5. Workers are not part of the product service …
6. Workers are emotionless and unaffected by factors such as pride, loyalty, and embarrassment.
7. Work is perfectly observable. Measurement error is ignored. No consideration is given to the possibility that observation changes performance (Hawthorne effect).

**The information in this table is reproduced from Boudreau et al.*(2003),page 183.*

Human fatigue which can be both physical and mental, is often blamed for errors leading to irreparable damage and near fatal accidents. The phenomenon of human fatigue and its description has led to an extensive research on the work-rest scheduling problem, and the job rotation scheduling problem (Finkelman, 1994; Rosekind et al., 1994; Westfall, 1996). It is one of the fundamental factors that has to be considered when assessing human performance.

Human motivation is also important in assessing human performance. Hertzberg et al. (1959), Tiejen and Myers (1998) and Eskildsen et al (2004) have shown that increased motivation leads to improved job performance and higher job satisfaction. Workers with higher motivation will tend to perform better. Hence, a person's motivation may affect job performance and should be considered in scheduling human task.

Stress is a factor that is known to impair attention, memory and action thereby leading to reduced human productivity. However, the impact of stress is not always negative. Measuring stress is the definitive resource for health and social scientists interested in assessing stress in humans. Reflecting the diversity of theoretical conceptions of stress, measuring stress provides integrative, incisive guidelines that will prove invaluable to students, clinicians and research in health and social psychology, medicine, nursing, epidemiology, sociology, and psychiatry.

Typical stressors include environmental stressors such as noise and vibration and psychological ones such as anxiety and fatigue (Cox and Griffiths, 1994).

Workload, mental and physical, can heavily influence human performance. Physical workload is defined as the ratio of the time required to complete a series of tasks to the time available in which to complete them. Mental workload is more generally defined as the ratio of the resources required to complete a series of tasks to resources available to complete as series of tasks (Wickens et al., 2004). The more common workload measures include task measures such as task performance measures, and physiological measures such as heart rate variability. High workloads often cause human fatigue and stress and can lead to disastrous results.

Learning and forgetting of skills can have a direct impact on human performance, e.g., in time and number of errors. When a person is acquiring a new skill, there is usually a learning curve, which quantifies the gain in efficiency as the person gains more experience performing a task or set of tasks (Neves and Anderson, 1981). Since learning and forgetting have an impact on productivity, include these factors in human task scheduling is relevant.

2.1.2.3 Empirical studies in workforce scheduling

Human considerations are typically observed in workforce scheduling problems that involve rotating workers among day, evening, and night shifts. Research on human task sequencing problems have been addressed using empirical methods, they demonstrate that strategic task sequencing positively impacts individual and hence, organizational performance by reducing fatigue, physical and mental stress, and risk associated with injury. Snook (1978) addressed the balance between productivity demands and safety concerns of the personnel involved in meeting these demands, showing that the profit made by the operation can be substantially diminished due to increased operating cost incurred by worker injury. Moray et al. (1991) introduces an alternative approach, using scheduling theory as a decision aid for human task scheduling. A more detailed framework of the analysis of strategic behavior using scheduling theory was presented by Dessouky et al. (1995). That work suggests that the analytic framework of classical scheduling theory is more appropriate to not only complement but also enhance traditional human factors approaches for studying strategic behavior. However, the framework makes no distinction between human and inanimate resources and does not account for human characteristics, handling workforce in scheduling as any limited resources. Researchers have

investigated the impact of various shift work scheduling strategies on the natural fluctuation of physiological human conditions that are governed by the Earth's day-night cycle (Tepas et al., 1997; Wickens et al., 2004a). Shift work is known to disrupt worker's natural circadian rhythms resulting in negative human conditions (Kostreva et al., 2002). Therefore, the primary objective of existing empirical approaches to workforce scheduling is to minimize disruption of circadian rhythms and the associated adverse effects. The effectiveness of a workforce scheduling strategy is measured by both human productivity and health. The literature's examination reveals that most empirical research involves subjecting different groups of employees to various scheduling strategies, collecting performance data through surveys and/or physiological measurements, and abstracting recommended strategies for developing effective shift schedules based on analysis of the collected data (Lac and Chamoux, 2004). Decision parameters of shift work scheduling strategies include speed of rotation, direction of rotation, shift duration, start time of the morning shift, and distribution of days off (Czeisler et al., 1982; Knauth, 1993; Monk, 2000).

2.2.Learning and forgetting

Worker-to-task assignment problem has been considered in the literature however researchers typically assume that workers perform at one productivity level or at a discrete set of productivity levels. Human have performance characteristics that differentiate them from machines. While a machine will always perform at a steady pace, the human performance displays alternate regimes of increase and decrease of productivity. Learning and forgetting patterns have been speculated to explain the existence of these regimes.

2.2.1. Learning

Learning has been discussed to take place in fourth ways; training, transfer of knowledge from organizational knowledge, learning by doing, and transfer of knowledge from other individuals.

2.2.1.1.Training

Training is an integral part of the job where the employees learn new skills while continuing to impact service to the organization, further the learning from training eventually result either from 'learning by doing' or from 'transfer of knowledge'.

2.2.1.2.Transfer of knowledge from organizational knowledge

Knowledge transfer in organizations is the process through which one unit is affected by the experience of another. Although it involves transfer at the individual level, the problem of organizational knowledge transfer transcends the individual level to consider transfer at higher levels of analysis, such as the team, product unit, or department. Changes in the knowledge or in the performance of the recipient units are the usual manifestations of occurrences of knowledge transfer in organizations. Knowledge transfer can then be quantified by measuring changes in knowledge or changes in performance. Darr, Argote and Epple (1995) develop a performance-based approach for measuring knowledge in the fast food stores and to investigate the extent to which productivity at a store was affected by the experience of the other stores in their franchise. Baum and Ingram (1998) analyzed the extent to which survival of hotels was affected by the experience in other hotels of the same chain. Assessing transfer through measuring changes in performance presents some challenges. Argote (1999) discussed the importance to control for factors in addition to the experience of other units that can potentially affect the performance of the recipient unit.

2.2.1.3.Learning by doing

We gain experience each time we repeat an activity. The experience can take the form of either or both improved manipulative skills and in process procedures. This is particularly significant during early ‘cycles’ or ‘repetitions’, each repetition adding both confidence and usually leading to an improved speed in accomplishing the performed task. Improvements are seen after large numbers of additional cycles are completed, this characteristic is called the ‘Learning’ phenomenon. Learning is also referred to as ‘experience’. The learning curved is the plotted relationship between the times taken to complete some activity a certain number of times. It is used for describing labor learning at the level of the individual employee in opposition to firm learning often called ‘Progress Functions’ which represent the improvements that occur at the process and plan level, right up to the level of the entire firm. Early learning curve research was an outcome of pioneering investigations done by experimental psychologists who studied reaction times and percentage recall of simple tasks as well as more complex tasks. Learning curves were first used in the aircraft industry and later, were spread to other industries such as the shipping, oil and automotive industries. The earliest learning studies were done by experimental psychologists Ebbinghaus (1885) suggests that the longer a person studies, the longer the retention, this was confirmed by Snoddy (1926). The most important learning theory

work and industrial application was done by Wright (1936) on aircraft production. His model, generally referred to as the 'Power Curve' or 'Power Model', despite dealing with very large product has been applied by a number of researchers (Conway and Schultz, 1959; Konz, 1983; Baloff, 1966, 1971) and has been extended to a wide range of human tasks (Newell and Rosenbloom,1981; Arzi and Shtub,1997). Learning curves usually have three major proxy types. They are cumulative output (Spence (1981), Fudenberg and Tirole (1983) and Lieberman (1984), cumulative investments (Sheshinski, 1967) and time (Rapping, 1965; David, 1970; Stobaugh and Townsend, 1975). Research on the acquisition of skills can be separated into two groups: The most extensive one deal with the behavioral scientists; Van Cott and Kinkade (1972), Anzai and Simon (1979), Mazur and Hastie (1978). Their research methods mainly consisted of laboratory experiments dealing with the responses to various stimuli. Carlson (1961) and Dejong (1957) proposed several hypotheses on how industrial workers learn. They found that there was a decrease from cycle to cycle in the number of times the eyes have to be refocused. Crossman (1959) says that an unskilled worker faced with a new task tries out several methods and retains the most successful ones. Levy (1965) conducted a taxonomy study on learning, suggesting a classification of learning into three categories: autonomous learning (improvement due to on-the job learning or training), planned or induced learning (result of a firm managerial actions for increasing the productivity or reducing production cost), and random or exogenous learning (improvement due to factors beyond the firm's control or expectation). Table 2.2. provides an exhaustive survey of models that explains the surge of productivity by learning by doing.

2.2.1.4.Transfer of knowledge from other individuals

At the individual level, transfer of knowledge may take place when working closely with fellow workers who are doing the same or related job. The ability of individuals to use knowledge accumulated by their colleagues has been discussed to turns on the rate of knowledge transfer inside the organization. Researchers who study experience working together suggest two distinct explanations for why working together improves performance. While the two explanations emphasize different mechanisms responsible for improved coordination, they highlight the ability of individuals to coordinate their activity. When studying the group dynamics, the importance of knowing "who knows what" has been underlined. Faraj and Sproull (2000), Larson et al. (1996), Lewis (2003), Moreland et al. (1996), and Wegner (1986, 1995) suggest that teams composed of individuals who have experience working together have more accurate

and shared sense of who knows what on the team. Liang et al. (1995) discuss the importance of knowing who knows what. They argue that it promotes the development of a more effective division of labor. Researchers studying experience working together in the market context have emphasized the importance of transaction partners learning how to govern their interaction. Uzzi (1996) suggest that multiple exchanges increase the likelihood of trust, Uzzi and Lancaster (2003) extend the idea by discussing how trust can promote the exchange of "private" knowledge and information. The two explanations, for why experience working together improves coordination, complement each other. Knowing who knows what is critical for defining roles and responsibilities inside the team and for assigning the most competent person to each role. However, this is only one half of effective coordination. The other half turns on the development of relationship-specific heuristics that enhance how well people performing distinct roles interact with each other. One can note the lack of model at the individual level that quantifies the amount of knowledge transferred and the impact of knowledge transfer in workforce scheduling.

2.2.2. Forgetting

Forgetting has been recognized as an important problem for production, Keachie and Fontana (1966) pointed out its significance within intermittent production. Baloff (1970) revealed the dramatic loss of productivity for short, intermittent runs due to loss of learning. Badiru (1994) shows, in an intermittent manufacturing, that learning and forgetting occur after one another supporting the concept of intermittent forgetting also referred to as the learn-forget-relearn cycle (Dar-El, 2000). Nembhard (2001) described the learning as a phenomenon that is observed by the decrease in production times per unit as operators gain experience by producing additional units. Forgetting takes place once there is a break of sufficient length. The production time of the first unit after returning from a break tends to be longer than the production time of the last unit produced before the break. Researchers have investigated factors effecting Forgetting. The break length appears to be the most common factor resulting to forgetting; the longer the break, the greater the amount of forgetting. Bailey (1989), Shtub, Levin and Globerson (1993) noted that the amount forgotten during an interruption is a function of the amount learnt so far and the length of the interruption. The nature of the task is also a factor effecting forgetting; the intensity of forgetting being greater for cognitive tasks than motor tasks, and greater for procedural task than declarative tasks (Bailey, 1989; Arzi and Shtub, 1997; Dar-El, 2000; Globerson et al, 1998). Models have been developed to estimate the amount of forgetting. These methods try to predict

the production time of the first unit following an interruption in which forgetting has occurred. Carson and Rowe (1976) proposed a learning and forgetting model in which forgetting is modeled by a power function similar to the learning curve. Nembhard and Osothsilp (2001) present a comparative survey of several forgetting models, the models are compared based on the measure of the amount of forgetting following a break in an environment where learning and forgetting occur intermittently. Table 2.3 provides a survey of forgetting models.

2.3.Learning and forgetting, scheduling

As mentioned on table 2.1, one of the general assumptions made in theory of scheduling is that workers are "stationary", no learning or forgetting effects, and they are considered to be identical in the sense that they work at the same speed. As pointed out earlier, the gap between theory and practice in scheduling has been acknowledged and has been discussed by a number of researchers. The lack of fit between the human aptitude and the formal or systemized element found in the scheduling methods and systems has been pointed to be one of the factors of the gap. Researchers typically assume workers operate at one productivity level. However, the reality is different, studies have proven that workers learn and forget at different rate. The remainder of this thesis focuses on human learning and forgetting aspects and their incorporation into workforce scheduling.

Scheduling models involving learning considerations are characterized by reduced processing times when tasks are performed repeatedly. Scheduling models with resource learning effects seem to have originated with a model that considers flow time minimization in a parallel machine environment (Meilijson and Tamir, 1984). Since, models and methods have been developed to address scheduling problems including single machine environment with sequence-dependent processing times (Biskup, 1999).

Table 2. 2. Learning Models (reported in Teyarachakul et al., 2011)

Model name	Functional Form	Parameters/variables
The cumulative Average Power Model by Hancock (1967).	$T(x) = T(1)x^{-m}$	$T(x)$ is the cumulative average unit time for completing x units.
Stanford's B Model.	$T(x) = T(1)(x + B)^{-m}$	B experience available for the program start
Dejong's learning Model (1957)	$T(x) = T(1)\{M + (1 - M)x^{-m}\}$ $0 \leq M \leq 1$	M time proportion that is incompressible
The Dar-El et al. Dual Phase Model (1995)	$T(x) = T_c(1)x^{-m_c} + T_\mu(1)x^{-m_\mu}$	$T_c(1)$ first unit time for the pure cognitive process $T_\mu(1)$ the first unit production-time for the pure motor process $-m_c$ learning slope for pure cognitive process $-m_\mu$ learning slope for pure motor process
The time Constant Model by Bevis et al. (1970)	$R_t = R_a + R_f\{1 - e^{-(t/\tau)}\}$	R_t the production rate at time t, R_a starting production rate that equals $\frac{1}{T(1)}$, R_f the steady state production rate during the transient part of the curve, and τ the time required to reach 63% of R_f,
The Exponential Model by Bevis et Al. (1970)	$T(x) = A * e^{-xB}$	A and B are the learning parameters
The Linear Model by Globerson (1980) and Hancock (1967)	$T(x) = A - B * x$	A and B are the learning parameters
The Hyperbolic Learning Function by Shafer et al. (2001)	$R_x = k\left\{\frac{x+p}{x+p+r}\right\}$	R_x production rate for x units of the cumulative work k the asymptotic steady state production rate, p the prior expertise, and r cumulative production needed to reach k/2

Table 2. 3. Forgetting Models (reported in Teyarachakul et al , 2011)

Model name	Functional Form	Parameters/variables
Power Models with Common experience on the forgetting curve Expression; VRIF, VRVF, LFCM	$T(x) = T(1)x^{-f}$	$T(x)$ the time for the xth unit of lost $T(1)$ the equivalent time for the first unit on the forgetting curve x the number of units that would accumulate if the break did not occur, and f is the slope of the forgetting curve
The Power Forgetting function by Globerson et al. (1989)	$F(T(S), I) = C_0 T(S)^{C_1} I^{C_2}$	C_0, C_1 and C_2 are parameters of the model
The Power Forgetting function by Bailey (1989)	$F(T(S), I) = C_1\big(T(1) - T(S)\big)\log(I) + T(S) + C_0$	C_0, C_1 are the forgetting parameters
The Single variable Sin function By Ramos and Chen (1994)	$Y(t) = ct^{-a}\frac{1}{2}(1 + \sin\ bt) + k$	t the time elapsed since worker got involved performing the task $Y(t)$ the performance level expected at time t a the learning rate, b the frequency of retraining, c the initial level of performance and k the desired long term performance standard
The Recency Model by Nembhard And Uzumeri (2000)	$T_f = T(1)\big(sR_s{}^{\beta}\big)^{-m}$ $0 \le R_s \le 1$	T_f unit output time after an interruption $R_s = 2\frac{\sum_{i=1}^{s}(t_i - t_0)}{s(t_s - t_0)}$ the recency of experiential learning, s the cumulative number of units produced, t_i the timestamp at the start of unit i, t_0 the earliest timestamp for the the worker, and β the degree to which the individual forgets the tasks.
Forgetting Models by Chiu et al. (2003)	$a_{j+1} = a_j\big(q_j + 1\big)^{-m}$ $+f\big(t_{j+1}\big)\Big(a_j - a_j\big(q_j + 1\big)^{-m}\Big)$ $f\big(t_{j+1}\big) = 1 - (1 - f(t_0))e^{-ct_{j+1}}$	a_j the first unit production cost of the jth lot, and $f\big(t_{j+1}\big)$ the forgetting rate if the interruption duration is t_{j+1} periods between lot j and lot j+1 and c the forgetting parameter

Pioneering work in scheduling models with sequence dependent processing times involving task deterioration are attribute to Gupta and Gupta (1988), Browne and Yechiali (1990). Voutsinas and Pappis (2002) review the modest number of instance involving two-resource flow shop, m-resource flow shop, and a parallel resource environment. Glazebrook and Mitchell (2002)

modeled deteriorating and improving jobs using Markov decision processes. These approaches, although each with benefits in its own right, have not directly consider the learning or forgetting effect at the human level.

The research literature on worker-to-task assignment, typically assume that workers operate at a steady productivity level or at a set of discrete productivity levels. However, workers learn and forget tasks based on how often they are assigned to a task and the lapse of time separating from the last assignment.

The necessity of incorporating worker learning and relearning behavior has been noted by Wisner and Pearson (1993). Their study illustrated that worker-relearning behavior has a significant impact on labor assignment policy, especially when a worker is required to transfer between tasks regularly. Nembhard and Norman (2007) remarked the scarcity of the literature on the assignment of individual workers to tasks based on learning and forgetting behavioral characteristics. They partially explained the little attention in the literature by the difficulty in collecting detailed performance data at the individual level. Moreover, a potentially significant limitation of many learning/forgetting models in the literature is that they commonly are designed as measurement tools and may not lend themselves well to use within other analytical or optimization frameworks. In fact, when individual worker learning and forgetting is incorporated, this introduces nonlinearity in the objective function.

Corominas et al. (2010) are the first to study the task assignment problem when work performance depends on experience of all tasks involved, the nonlinearity of their performance function, thus of the scheduling problem, was overcome by using a linearization technique. In their work the main objective is to minimize completion times. They do not consider an explicit model of learning. Nembhard (2001) developed a heuristic approach for assigning workers to tasks based on individual learning rates. The heuristic developed significantly improved overall productivity under empirically observed conditions and under many experimental conditions. Nembhard and Norman (2007) proposed a mathematical formulation for assigning worker to station in independent jobs dependent stations systems (referred to as serial system or flow line). As far as the authors are aware, it is the first and explicit used of a learning model in a mathematical formulation of a worker to task assignment problem.

CHAPTER 3

PARALLEL SYSTEM SCHEDULING WITH HUMAN PERFORMANCE VARIABILITY

3.1. Introduction

Workforce scheduling has been studied from a variety of perspectives, in particular in relation to task management based on processing times, minimizing completion or latency times, or prioritization. Resolving such problems can be difficult, time intensive, or problematic. In recent years, the scheduling and assignment of human operators to tasks has emerged as an important field of study. This chapter focuses on scheduling workers to a set of parallel tasks, whereby both the stations and the jobs are independent of one another, taking into account the significant human capacity for learning and the tendency to forget (alternately referred to in the literature as improvement and retention, respectively). It is hoped that the results of the study may enable solving a wider range of scheduling problems and improve upon current methods and practices in this field. The remainder of this chapter is organized as follows. Section 2 presents a mathematical formulation of the scheduling problem in an environment with parallel tasks and learning, along with the resultant structure of the optimal solution. In Section 3, the result is generalized to other problem types by varying the ratio of workers to tasks, the degree of improvement and retention, dependent and stochastic learning, and demand requirements.

3.2. Model

A straightforward formulation of the multi-period parallel system scheduling problem is presented, with the aim to maximize the overall system output. It is assumed that worker-productivity is specific to the task and is a function of the skills the worker possesses and the cumulative number of previous assignments on that particular task. It is assumed that two workers cannot perform a task simultaneously and tasks are performed independently of one another. The model assumptions are described in detail in section 2.2.1, followed by the formulation of the model presented in section 2.3.2, section 2.3.3 is devoted to the structure of the optimal solution.

3.2.1. Assumptions

The mathematical programming model developed in this study is based on the following assumptions:

- The time horizon is divided into periods of the same length. In each period, a worker is assigned to a single task or to none, whereby changes in assignments during a given period are not possible. This setting has ties to practical applications, and the advantage of facilitating the formalization of the problem.
- The expected worker performance on a task is obtained as a function of the sum of the experiences on the task.
- A task is performed at a unique station and a station can be used to perform only one type of task. That is, task or station will be either used to define the same entity.
- Each station has an infinite input buffer; hence, a worker assigned to a station is never starved for work.
- A worker can perform at most one task, and a task can be performed by at most one worker, there is a one-to-one relationship between workers and tasks; i.e. workers cannot multitask and tasks cannot be shared. Assignments are binary, thus, in a given period a task can only be assigned to one worker or none.

3.2.2. Model Formulation

In this section, a formal description of the model examined in this chapter is presented following some preliminary notation.

J Number of tasks $j = 1, \ldots, J$

I Number of workers $i = 1, \ldots, I$

T Number of periods $t = 1, \ldots, T$

$x_{i,j,t}$ Binary variable that indicates whether task j is performed by worker i during time t, where $i = 1, \ldots, I, j = 1, \ldots, J, t = 1, \ldots, T$.

$O_{i,j,t}$ The output of worker i performing task j during period t. $O_{i,j,t}$ is a function of worker skill, and of previous assignments (experience) on task j.

f_{ij} Estimates the productivity of worker i on task j. It is considered that the performance model is a differentiable non-decreasing function of the cumulative time the worker spent on the task.

The worker-task assignment problem with multi periods of production is formulated as in problem $\boldsymbol{P}$ (Equations 1-4).

$$\boldsymbol{P} = \begin{cases} max \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{t=1}^{T} \boldsymbol{x_{i,j,t}O_{i,j,t}} & (\boldsymbol{3-1}) \\ \sum_{j=1}^{J} \boldsymbol{x_{i,j,t}} \leq \boldsymbol{1} \quad \forall\, \boldsymbol{i}, \forall\, \boldsymbol{t} & (\boldsymbol{3-2}) \\ \sum_{i=1}^{I} \boldsymbol{x_{i,j,t}} \leq \boldsymbol{1} \quad \forall\, \boldsymbol{j}, \forall\, \boldsymbol{t} & (\boldsymbol{3-3}) \\ \boldsymbol{O_{i,j,t}} \quad \leq \quad f_{ij}\left(\sum_{k=1}^{t} \boldsymbol{x_{i,j,k}}\right) \times length\ of\ period \quad \forall\, \boldsymbol{i}, \forall\, \boldsymbol{j}, \forall\, \boldsymbol{t} & (\boldsymbol{3-4}) \end{cases}$$

In problem $\boldsymbol{P}$, Equation (3-1) is gives the expression for the total production in the system, whereby we sum over all tasks, all workers and all time periods, to obtain the system output. Equation (3-4) provides a rectangular approximation of the learning function in order to obtain $O_{i,j,t}$, the output of worker i on task j during time t. The error from this approximation is negligible for sufficiently short period lengths. Constraint (3-2) implies that a worker can perform no more than one task at a time, and Constraint (3-3), that each task may only be paired with one worker at a time.

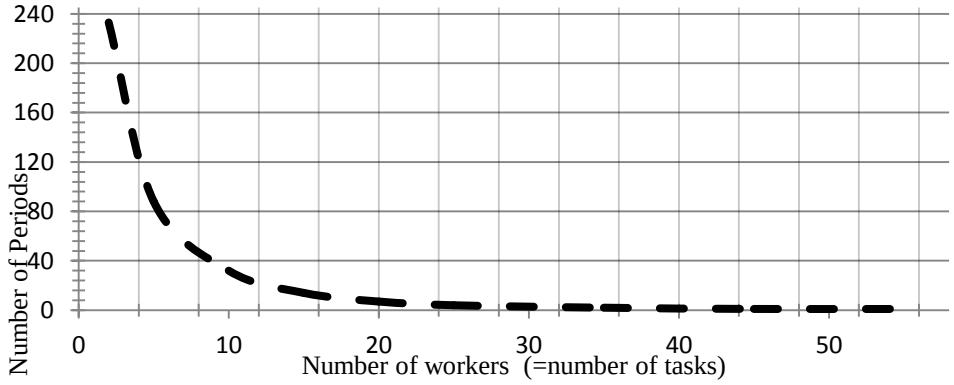


Fig 3. 1. Problem Instance Solvable Frontier for P (Gams-Baron)

Figure 3.1 illustrates the practical limitations of formulation $\boldsymbol{P}$ with respect to the number of workers, tasks and time periods that can be solved (*GAMS* software with *Baron* solver*)*. The curve represents the largest solvable problem instances when formulation $\boldsymbol{P}$ is implemented with

a heterogeneous workforce. For example, when the number of workers and tasks is *10*, the maximum number of periods that can be considered and effectively solved does not exceed *40*. These practical limitations are an important motivation of the current study.

If $\mathcal{A}(I,J)$ is the set of all possible assignments for I workers and J tasks, the size of $\mathcal{A}(I,J)$ is $\frac{J!}{(J-I)!}$ assuming $(J \geq I)$. A schedule S for a production time horizon, T is a sequence of possible assignments and can be represented as $S = A_1 \rightarrow \cdots \rightarrow A_T$, where A_q is a feasible assignment for the I workers at period t, $1 \leq t \leq T$. Thus, a feasible assignment A_q is a matrix of size $I \times J$ with coefficients defined as follows:

$$(A_q)_{i,j} = \begin{cases} 1, if\ worker\ i\ \ performed\ \ task\ j \\ in\ assignment\ \ A \\ 0, otherwise. \end{cases}$$

For T periods of production, the size of the set of possible schedules is $\left[\frac{J!}{(J-I)!}\right]^T$. Hence, for a large value of T, or J, a direct enumeration of the solution space is computationally prohibitive. Thus, reducing the number of feasible schedules considered can significantly improve the maximum size problem instance that can be solved, as well as improve the computational time efficiency. In the following subsection, we analytically derive the structure of the optimal solution to problem ***P***, which in turn yields a problem formulation with reduced dimension.

3.2.3. Structure of the optimal schedule

Theorem 3.1 When maximizing total production, with the number of tasks equal to the number of workers and a twice differentiable monotonically non-decreasing performance model of the time, the optimal schedule is a schedule with specialized workers. That is, no change is required within the task assignment problem and workers are assigned to a single task.

Proof: In Appendix A

Theorem 3.1 states that when the objective is to maximize the production of a system of independent stations and independent jobs, the formulation ***P*** can be converted without loss of optimality to a new formulation ***P′*** in which, $\boldsymbol{F_{i,j,t}}$ is the production of worker *i* through period *t*

at task *j*. The Objective function for ***P'*** follows equation $(3-5)$, with constraints, as before in Equations $(3-2)-(3-4)$.

$$max \sum_{i=1}^{I} \sum_{j=1}^{J} \boldsymbol{x_{i,j} F_{i,j,T}} \qquad (3-5)$$

The complexity of solving the worker-task scheduling problem with learning, in a parallel system, thus depends only on the number of tasks and number of workers. This implies that one can solve formulation ***P*** via formulation ***P′*** by initially computing $\boldsymbol{F_{i,j,T}}$ (i.e., $\boldsymbol{F_{i,j,T}} = \int_0^T f_{ij}(t)dt$, $f_{i,j}$ the productivity rate of worker *i* on task *j*). Despite the problem dimension reduction, solving ***P'*** provides exact solutions to the original problem ***P*** based on Theorem 3.1.

Theorem 3.1 stipulates that specialization is an optimal policy when maximizing production. However in some special circumstances, despite the apparent facility of implementing specialization, switching workers may provide equivalent production and some ancillary benefits. This is possible if tasks are similar and knowledge/skill acquisition can be transferred from one task to another. That is, in some special cases, learning on one task, may provide sufficient expertise on other tasks to provide an equivalent level of performance despite switching tasks. Thus, there may be secondary benefits in terms of increased workforce flexibility as well as the potential benefits of balancing production.

3.2.4. Computational Comparisons

This section examines some practical numerical characteristics of using solutions from problem formulation ***P'***. These include computation time comparisons solving ***P*** directly, and an investigation of the relative impacts of problem dimensions on the time complexity when solving ***P'***.

3.2.4.1. Computational Time of *P* and *P'*

Some computational comparisons are conducted using the three-parameter hyperbolic learning model described in Mazur and Hastie [29]. However, the result in Theorem 3.1 is general and not dependent on any specific learning model. In the hyperbolic model, the productivity of worker *i* on task *j* during time *t*, is:

$$f_{ij}(t) = k_{i,j}\frac{t + p_{i,j}}{t + p_{i,j} + r_{i,j}} \qquad (3-6)$$

where, $k_{i,j}$ is the asymptotic constant learning, $p_{i,j}$ is the initial productivity level (i.e., for the first unit work), and $r_{i,j}$ is the learning parameter.

The potential usefulness and effectiveness of using the result in Theorem 3.1 is illustrated by solving a range of workforce/task size problems, the results of which are displayed in Table 3.1, whereby the parameters of the workers are generated via simulation. A multivariate normal (MVN) distribution is fitted to an empirical population of worker characteristics, *k, p, r* parameters, and subsequently sample from this MVN distribution in the simulations. Characteristics of the MVN distribution (mean vector and variance-covariance matrix) are based on empirical data adapted from the study by Shafer et al. (2001), where inspection time data from 176,000 items produced by 75 individual workers was collected over a 6-month period at a consumer electronics plant. The respective vector of mean and variance-covariance matrix of [*ln(k), ln(p), ln(r)*] are approximated by $\boldsymbol{\mu}, \boldsymbol{\Sigma}$ as:

$$\boldsymbol{\mu} = [3.36 \quad 4.66 \quad 4.83] \qquad \boldsymbol{\Sigma} = \begin{bmatrix} 0.054 & 0.397 & 0.374 \\ 0.397 & 7.821 & 5.430 \\ 0.374 & 5.430 & 4.923 \end{bmatrix}$$

where there is significant correlation (at 5%) between all pairs of parameters. Moreover, $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are derived directly from the empirical data. The Ryan-Joiner test for normality (Ryan and Joyner, 1973) indicates that these data are not significantly different from normal (at 5%).

Table 3. 1. Summary CPU Times for Sample Problem Instances

Problem Size	**Solution to P**	**Solution to P'**
(*I*/*J*/*T*)	**CPU time (sec)**	**CPU time (sec)**
5/5/7	2.49 (0.11)	0.17 (0.04)
8/8/9	32.42 (2.44)	0.17 (0.01)
11/11/12	361.45 (95.35)	1.25 (0.95)

[a] Average (standard deviation) over 20 replications for each problem size

The results in Table 3.1 present 3 problems, each described by the triplet *I/ J/ T (*number of workers/ number of tasks/ number of periods). For each problem, 20 replications were generated, wherein the *k/p/r* parameters for a worker assigned to each task are sampled from the MVN distribution. The performance is compared to that of the straightforward approach, which consists of solving formulation ***P*** (*Gams* – BARON, MINLP solver) to the performance of when solving ***P'*** (*Gams* - Cplex LP solver). For each of the 3×20 runs, the two approaches generate the same solutions with respect to the system output. However, as Table 3.1 illustrates, ***P*** and ***P'*** differ significantly with respect to computation times, and this difference is much more pronounced as the instance size increases.

3.2.4.2. Computational Time for Solving P'

This section studies potential impacts of the number of workers, the number of time periods, and the heterogeneity among worker learning characteristics on the solution time. There are a total of 27 treatments in the comparisons (3 levels of *I*) $\times$ (3 levels of *T)* $\times$ (3 levels of δ), with 20 replications for each treatment. Worker characteristics are sampled from the MVN distribution using $\delta \times \mathbf{\Sigma}$, where the heterogeneity, δ, is a controlled factor and $\mathbf{\Sigma}$ is the variance-covariance matrix of the worker parameters, as before.

The results summarized in Table 3.2, and figure 3.2, indicate that the number of workers and the number of periods are significantly positively related to the run time when solving ***P'***, whereas worker heterogeneity is not significant. That is, performance is not dependent on the workers' characteristic parameters. The numerical examples reveal the limits of the straightforward approach when solving the worker-task assignment with learning (as illustrated in figure 3.1). In contrast, the size of problems that can be solved using ***P'*** is ultimately determined by the limits of the LP solver employed. In the current study, the solver could handle 96 workers/96 tasks/246 periods, which we note is a substantial improvement over the straightforward approach (solving ***P*** directly).

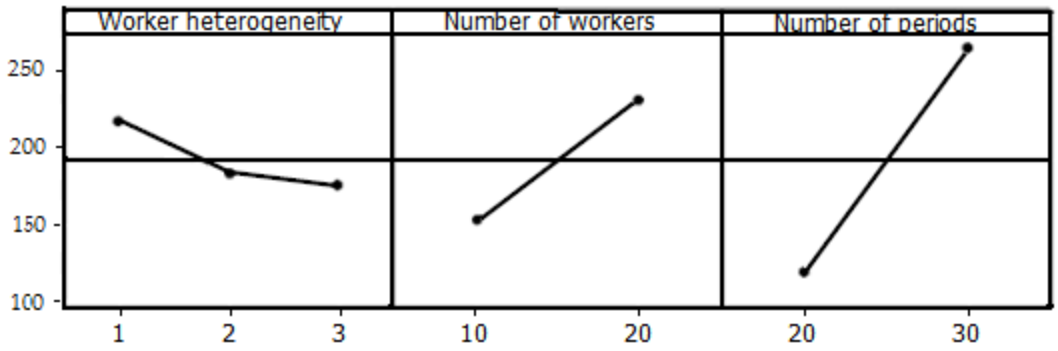


Fig 3. 2. Main effects for P' Computational time (seconds)

Table 3. 2. Summary of ANOVA for Computational time solving P'

Source		*F*	*p*
Number of workers	I	11960.45	0.000
Number of periods	T	505.02	0.000
Worker Heterogeneity	δ	0.02	0.983

3.3.Extensions

In this section, the result in Theorem 1 is extended to several important and more general conditions, including consideration of cases where $I \neq J$, dependent learning, stochastic learning, learning and forgetting, and demand requirements. These individually are introduced as independent generalizations of Theorem 1.

3.3.1. Number of workers lesser than the number of tasks $\boldsymbol{I < J}$

Theorem 1 addresses the balanced case of $I = J$. However, in general practice it is common for systems to have either greater or fewer tasks than workers. The case where $I < J$ is tackled

The number of workers is fewer than the number of tasks, let $N = J - I$ denotes the number of *surplus tasks* compared to the scenario in Theorem 1. Thus, $\tilde{I} = I + N$ is equal to the number of

tasks *J*, where workers [1 .. *I*] represent actual workers, and workers $[I+1\,..\,J]$ represent an additional requirement, termed *artificial* workers.

The set of skills is defined as follows:

For $i \leq I$, the initial distribution of worker's skills is maintained.

For $I < i \leq J$,

$$y_{i,j} = \varepsilon,\ for\ all\ i,j$$

In formulation ***P***, Equation 4 becomes,

$$\boldsymbol{O_{i,j,t}} \quad \leq \quad \varepsilon \times (\text{Period Length}) for\ i = I+1 \ldots J, \qquad j = 1 \ldots J, \qquad t = 1 \ldots T$$

3.3.2. Number of workers greater than the number of tasks $\boldsymbol{I > J}$

The number of workers is greater than the number of tasks, let $M = I - J$ denotes the number of *surplus workers*. Thus, using a similar approach as above, *M* artificial stations are introduced, such that the number of tasks $\tilde{J} = J + M$ is equal to the number of workers *I*, where the "real tasks" are *j* = [1 .. *J*] and the artificial tasks are $j = [J+1\,..\,I]$.

The set of skills is defined as follows:

For $j \leq J$, the initial distribution of worker's skills on task *j* is maintained.

For $J < j \leq I$,

$$y_{i,j} = \varepsilon,\ for\ all\ i,j$$

Thus, in formulation ***P***, Equation 4 becomes,

$$\boldsymbol{O_{i,j,t}} \quad \leq \quad \varepsilon \times (\text{Period Length}) \quad for\ i = 1 \ldots I, \qquad j = J+1 \ldots J, \qquad t = 1 \ldots T$$

ε is chosen to be small enough, so that the contribution of the artificial worker to any task will be negligible. A suitable choice would be such that $\varepsilon * T$ is negligible compared to *1*, where *T* is the number of periods available for task completion. Intuitively, when *I*< *J*, the *I* workers are allocated to the most productive of the *J* tasks, with the remaining tasks not performed. When *I* > *J*, scheduling workers to switch between tasks at any time is suboptimal. That is, workers should be assigned to a single task during the entire period to maintain optimal production. The result

also indicates that the *M* surplus workers will not be used during the time horizon for the current schedule.

3.3.3. Tasks with Dependent Learning (Task Similarities)

Olivella (2007) examined task performance, considering experience gained by performing similar and identical tasks, as well as how recent the experience was, based on a modification to the learning and forgetting model described by Nembhard and Uzumeri (2000). The result in Theorem 1 is extended, to address general learning when performance on a task can be influenced by assignment to other similar tasks. For each task *j*, a set of tasks T_j that are similar to task *j* is defined. In Equation (3-4), the productivity function of worker *i* on task *j* at time *t*, $f_{ij}(t)$ depends on the worker skills on task *j* and the cumulative experience of worker *i* on task *j*, $\sum_{k=1}^{t} \boldsymbol{x_{i,j,k}}$. When task similarities exist, the productivity of worker *i* on task *j* is influenced by the worker's skills on task *j*, the cumulative worker experience on task *j*, $\sum_{k=1}^{t} \boldsymbol{x_{i,j,k}}$, *and* the cumulative experience on tasks similar to task *j*, $\sum_{q \,\epsilon\, T_j, q \neq j} {\alpha^j}_q \left(\sum_{k=1}^{t} \boldsymbol{x_{i,q,k}}\right)$. Where ${\alpha^j}_q$ is a real number representing the amount of acquired knowledge on a task *q*, $q \,\epsilon\, T_j$ that can be associated as knowledge for task *j*. For instance, $x_{i,q,k} = 1$ is equivalent to a learning of ${\alpha^j}_q \times 1$ for task *j*.

The following constraints on *similarity*, ${\alpha^j}_q$, for all *j* are added :

$$0 \leq {\alpha^j}_q \leq 1 \qquad (3-7)$$

the experience gained from a task similar to task *j*, e.g. $\boldsymbol{x_{i,q,k}}$ where $q \,\epsilon\, T_{j,}$ will never exceed the experience gained if the worker was assigned to the task *j* itself during the same amount of time.

$$\sum_{q \,\epsilon\, T_j, q \neq j} {\alpha^j}_q \leq 1 \qquad (3-8)$$

performing several tasks $q, q \,\epsilon\, T_j$, similar to a task *j* could not result in a learning for task *j* superior to the learning the worker would have reached if he or she was assigned to task *j* during a time equals to the sum of the duration of each assignment on these tasks *q*.

Thus, allowing $n_{ijt} = \sum_{k=1}^{t} x_{i,j,k}$ and using Equation $(3-2)$ and $(3-3)$ lead to:

$$\sum_{i=1}^{I} n_{ijt} \leq t \qquad i = 1 \dots I, \qquad j = 1 \dots J, \qquad t = 1 \dots T \tag{3-9}$$

$$\sum_{j=1}^{J} n_{ijt} \leq t \qquad i = 1 \dots I, \qquad j = 1 \dots J, \qquad t = 1 \dots T \tag{3-10}$$

Relaxing the condition that n_{ijt} is an integer, the derivative of f_{ij} with respect to $n_{ijt}, for\ all\ j$ can be written as:

$$\frac{\partial f_{ij}}{\partial n_{iqt}} = \begin{cases} \alpha^{j}{}_{q} \dfrac{\partial f_{ij}}{\partial n_{ijt}} & for\ all\ \ i,t \quad for\ q \in T_j \\ 0 \quad for\ all\ \ i,t \qquad for\ q\ \notin\ T_j \end{cases} \tag{3-11}$$

Given the constraint on α_q the following inequality is obtained

$$\sum_{q \neq j} \frac{\partial f_{ij}}{\partial n_{iqt}} \leq \frac{\partial f_{ij}}{\partial n_{ijt}} \tag{3-12}$$

Thus, Equation (3-12) implies that the knowledge gained by assigning worker *i* to task *j* during time interval *t* is greater than that gained by performing a task similar to task *j* during the same period. Moreover, the cumulative knowledge gained from performing a task similar to task *j*, is less than the learning the worker would have gained if s/he performed task *j*. Hence, when seeking to increase task productivity, assigning a worker to the task itself will ultimately be preferred to performing other similar tasks. Incorporating knowledge gained from these other tasks will not affect the structure of the optimal schedule when the objective is to maximize the total production.

3.3.4. Stochastic Learning

Theorem 3.1 assumes a deterministic learning model. This section extends the result to the case of stochastic learning, which has been proposed to better represent workplace environments (see e.g., Uzumeri and Nembhard, 1998).

Consider the following condition for a stochastic learning model:

$$y_{i,j}(s,\omega) \leq y_{i,j}(t,\omega) \text{ for all } s, t \text{ such that } 0 \leq s \leq t, \text{ and for all } \omega \in \Omega \tag{3-13}$$

where Ω is the set of all possible events, representing stochastic states, and $y_{i,j}$ is the productivity of worker *i* on task *j*. Equation (3-13) stipulates that by performing the same task during a longer period of time, a worker can only increase their learning on a given task. By applying a stochastic model of learning, the multi-period worker-task scheduling problem can be formulated by replacing the objective function by that in Equation (3-14):

$$max \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{t=1}^{T} \boldsymbol{x_{i,j,t}} \mathbb{E}[\boldsymbol{O_{i,j,t}}] \qquad (3-14)$$

The objective in Equation $(3-14)$ is to maximize the expected production, where now $\boldsymbol{O_{i,j,t}}$ is a stochastic process that estimates the production of worker *i* on task *j* at period *t*, and $\mathbb{E}[\boldsymbol{O_{i,j,t}}]$ is the expected production**.** When considering a stochastic model of productivity as in Equation $(3-13)$, Theorem 1 holds and specialization of the workforce remains optimal. The proof of Theorem 1 relies on the assumption that the learning curve of the worker on a given task is an incremental function of the time s/he spent on a task. Thus, Equation $(3-13)$ is the stochastic version of the assumption that the rate of increase is non-deterministic. Replacing the terms related to the production by expected production, yields the optimal schedule that maximizes the expected production, and each worker is assigned to a single task. The stochastic task-worker assignment problem under condition $(3-13)$ can be formulated using the objective function in Equation $(3-13)$, where $\mathbb{E}\left[\boldsymbol{P_{i,j,T}}\right]$ is the expected production for the entire time horizon *T*,

$$max \sum_{i=1}^{I} \sum_{j=1}^{J} \boldsymbol{x_{i,j}} \mathbb{E}\left[\boldsymbol{P_{i,j,T}}\right] \qquad (3-15)$$

3.3.5. Learning and Forgetting Models

Since models of productivity that incorporate learning and forgetting were introduced to represent tradeoffs in manufacturing and service environments, Teyarachakul, Chand and Ward (2011) proposed some basic properties of forgetting functions, suggesting that the process of forgetting is an incremental function of the length and the number of the interruptions, whereby there is more forgetting if there is more to forget. A first illustration is conducted using a specific

learning/forgetting function, (see Thomas and Nembhard, 2004), whereby the productivity rate of worker *i* on task *j* during time period *t* is given by O_{ijt}:

$$\begin{cases} O_{ijt} = l_{ij} + K_{ij}L'_{ijt}F'_{ijt} \\ L'_{ijt} = 1 - exp\left(-\frac{1}{L_{ij}}\sum_{k=1}^{t} x_{i,j,k}\right) \\ F'_{ijt} = exp\left[\frac{1}{F_{ij}}\left(\sum_{k=1}^{t} x_{i,j,k} - t\right)\right] \end{cases} \qquad (2-16)$$

l_{ij} is the initial expertise for worker *i* on task *j*.

K_{ij} is the steady-state productivity rate for worker *i* on task *j*.

L_{ij} is the learning rate for worker *i* on task *j*.

F_{ij} is the forgetting rate for worker *i* on task *j*.

L'_{ijt} can be viewed as the learning contribution to the productivity rate function, whereas F'_{ijt} accounts for the forgetting effect.

Since $\sum_{k=1}^{t} x_{i,j,k} \leq t$ and $F_{ij} > 0$, the following inequality holds

$$O_{ijt} \leq l_{ij} + K_{ij}L'_{ijt} \quad i = 1 \dots I, \quad j = 1 \dots J, \quad t = 1 \dots T \qquad (2-17)$$

That is, the expression $l_{ij} + K_{ij}L'_{ijt}$ satisfies the properties of a general learning function (Badiru, 1992). Thus, solving formulation ***P*** with the productivity rate given by $l_{ij} + K_{ij}L'_{ijt}$ provides an upper bound. Theorem 1 stipulates that solving formulation ***P*** is equivalent to solving the assignment problem with specialized workers. Moreover, when considering a schedule of specialized workers, O_{ijt} becomes the upper bound $l_{ij} + K_{ij}L'_{ijt}$. Thus, as the optimal schedule with specialized workers maximizes at upper bound, it is also optimal when forgetting occurs if the workers are not acquiring additional practice, such as when switching among tasks. Hence, the structure of the optimal schedule with both learning and forgetting remains the same as that for learning alone and specialization is optimal.

3.4.Scheduling with Demand Requirements

In formulation ***P***, the structure of the optimal schedule in a system with independent stations and independent jobs was established without any production requirements on the tasks. In this section, the case of a specific demand associated with each task is investigated. The formulation of this problem, $\boldsymbol{P_D}$, is as before, in Equations $(3-1) - (3-4)$, with the addition of the set of demand constraints in Equation $(2-18)$. This new constraint stipulates that the total production on task j, meets or exceeds the demand requirement for the associated task D_j.

$$\sum_{i=1}^{I}\sum_{t=1}^{T} x_{i,j,t} O_{i,j,t} \geq D_j \quad \forall\, j \qquad (3-18)$$

These additional constraints imply that the optimal solution structure of Theorem 1 does not necessarily hold, and is straightforward to verify. For example, a highly capable worker may be required to perform multiple tasks in order to meet the minimum production quantities demanded, thereby foregoing total aggregate output, in favor of meeting all task requirements. It is nonetheless possible to simplify formulation $\boldsymbol{P_D}$ based on Theorem 3.2, which allows for a reduction in the complexity of the problem to be solved. The result is based on considering *stint*; a contiguous time interval where a worker is assigned to a particular task.

Theorem 3.2 Within parallel systems, when the demand is set for production during the entire time horizon and considering a differentiable non-decreasing model of performance, the maximum number of stints per worker per task is one.

Proof : In appendix B.

Based on the result of Theorem 3.2, problem $\boldsymbol{P_D}$ may be reformulated as in problem formulation $\boldsymbol{P'_D}$ as follows:

$$max \sum_{k}^{T}\sum_{i=1}^{I}\sum_{j=1}^{J} M_{i,j,k} \times y_{i,j,k} \qquad (3-19)$$

$$\sum_{k=1}^{T}\sum_{j=1}^{J} k \times y_{i,j,k} \leq T \quad \forall\, i \qquad (3-20)$$

$$\sum_{k=1}^{T}\sum_{i=1}^{I} k \times y_{i,j,k} \leq T \quad \forall j \qquad (3-21)$$

$$\sum_{k=1}^{T} y_{i,j,k} \leq 1 \quad \forall i, \forall j \qquad (3-22)$$

$$\sum_{k=1}^{T}\sum_{i=1}^{I} M_{i,j,k} \times y_{i,j,k} \geq D_j \quad \forall j \qquad (3-23)$$

$y_{i,j,k}$ is an binary variable with k the length of the assignment of worker i on task j .

$M_{i,j,k}$ is the output of worker i on task j during the assignment of duration k. $\left(M_{\ i,j,k}\right)_{1\leq i\leq I, 1\leq j\leq J, 0\leq k\leq T}$ is a three-dimensional matrix of size $I \times J \times T$, which may be pre-calculated using the following expression $\boldsymbol{M_{i,j,k}} = \int_0^{\boldsymbol{k}} f_{ij}(v)dv$.

Equation $(3-19)$ describes the objective function which the sum over all workers, all tasks and all possible length of stint. Constraints $(3-20)$ and $(3-21)$ ensure that a task is not performed more than the time devoted to production and the time workload of a worker cannot exceed the production time. Constraint $(3-22)$ stipulates that for a given task, a worker has at maximum one assignment to that task. Constraint $(3-23)$ ensures that we satisfy the demand of the whole time horizon. Because it is possible that the workforce is incapable of meeting all demand constraints, infeasible problem instances are possible in this case, and alternate approaches (e.g., goal programming) may mitigate such issues.

Conclusions

The present study augments extant knowledge on optimal scheduling in parallel production systems with learning. To accomplish this, the structure of the optimal schedules is established, leading to a problem statement with reduced complexity. This substantively increases the maximum problem size that can be solved, as well as the corresponding computation time. It is specifically showed that the schedule for workers on a set of parallel tasks should be such that workers are specialized to tasks. The result holds for more, equal, or fewer numbers of workers relative to the number of tasks, with learning and/or forgetting, with tasks that have some similarities, and with stochastic processing times. In the case when there are production requirements on each task, workers are not necessarily specialized. However, they

will have no more than a single stint on any one task. That is, a worker may need to perform multiple tasks, but if so, they will complete all the required work for one task during the time horizon before starting any additional task(s). The possibility of cross training in this context is a less strong result than that of pure specialization. However, this result similarly allows for solving problems of reduced complexity.

Even though numerical illustrations were conducted based on the three-parameter hyperbolic model adopted from Mazur and Hastie (1978), the analytical results are not model specific, and are general in nature without considerations on the task complexity or the variance among the workforce. The result applies to any distribution of task complexity and worker learning / forgetting parameters. Also, the structure of the optimal schedule with the objective to maximize the total production is insensitive to the notion of time granularity in the schedule (see Barachini, 1994). The structure of the optimal schedule suggests that schedules with worker rotation will be suboptimal.

There are numerous appropriate rationales for cross training a workforce on a set of parallel tasks. These might include but are not limited to, increased job satisfaction, reduced fatigue, improved motivation, flexibility, and workforce retention. For instance Volpe et al. 1996, Ebeling and Lee (1994), Gerwin (1993), Molleman and Slomp (1999) suggest that cross-training increases workforce flexibility, which is a valuable asset in a highly uncertain and competitive environments. Nonetheless, any such rationale must entail some costs with respect to quantitative system output, since training team members to perform duties of other team members implies switching, and tradeoffs between the need to be able to respond to changes and the desire to maximize production. The optimal structures shown in this paper allow for quantifying these costs and tradeoffs, particularly for larger problem instances. Qin and Nembhard (2007) suggest a framework using real options to make cross-training decisions based on the tradeoffs between productivity, and future flexibility.

Secondly, the result described here applies to a set of parallel tasks, without process dependencies. A clear extension of this work is in the area of serial systems. The development of optimal strategies for serial systems is of ongoing concern in the literature. For example, Bartholdi and Eisenstein (1996) showed that, in production lines where workers are sequenced from the slowest to the fastest, a stable partitioning of work would eventually emerge. As a

result, an optimal corresponding production rate could be found. Armbruster, Gel and Murakami (2007) extended these results by incorporating worker learning. However, an important limitation of these studies is that the results apply to the steady state behavior, that is, once learning and other transients have subsided. Extending the understanding of workforce dynamics in serial systems, in particular toward shorter-term transient operations is of considerable interest for future research.

CHAPTER 4

SERIAL SYSTEM SCHEDULING WITH HUMAN PERFORMANCE VARIABILITY

4.1 Introduction

Following the study of parallel system, this chapter is devoted to the workforce scheduling problem in sequential *flow line* systems. A discussion on the impact of within-worker and between-worker variability is conducted as well as the selection of scheduling policies between fixed and work-sharing systems. The methodology includes human performance modeling with the objective to maximize throughput with general results with respect to productivity model. Determining and characterizing the optimal switching time between workers in work-sharing systems are of interest. A closed form solution is established in the case of two and three tasks with two workers. For a general number of workers and tasks, the maximum number of changes between assignments is established as well as a bound on throughput. Results allow one to solve workforce scheduling problems reduced complexity given the current set of bounds and conditions.

4.2 A Flow Line with Two Workers and Two Tasks

The consideration of a two-task flow-line will illustrate some useful insights and form a basis for considering larger systems. Nembhard and Norman (2004) suggested a math-programming model to schedule workers to tasks in serial system. Here, the main elements of their formulation are highlighted:

$$maximize \sum_{i=1}^{I}\sum_{t=1}^{T} O_{i,j,t}$$

s.t

$$O_{i,j,t} \leq X_{ijt}P_{ijt} * (period\ length) \qquad j = 2, \dots, J \;\; t = 2, \dots, T$$

$$B_{j,t} = B_{j,t-1} + \sum_{i=1}^{I} O_{i,j-1,t} \quad - \sum_{i=1}^{I} O_{i,j,t} \quad j = 2, \dots, J \;\; t = 2, \dots, T$$

The decision variable X_{ijt} defines the assignment of each worker at time *t*. The first constraint estimates the output of worker i on task j at time t while the second constraint defines the buffers' dynamics with their levels represented by variable $B_{j,t}$. Within the model, we have that:

- Time is separated into periods represented by an integer t that varies from 1 to T. A worker performs at most one task during each time period.
- Workers are heterogeneous with respect to initial productivity, learning, and forgetting rate for each task. Their productivity is measured by P_{ijt}.

Within the flow line, two opposing scenarios form bounds on practical operations: Instantaneous availability and batch transfer. That is, instantaneous availability occurs when the production planning is divided into working periods and work. Once work is completed at a station, it is immediately available to be processed at the subsequent downstream station during the same period. With batch transfers, production is also divided into working periods, but a job processed at a station is only available for work at the downstream station during the next period. These two scenarios constitute an approximation of real situations in flow line systems where an item can only be processed at the downstream station if it has already passed the upstream station. Conceptually, as period length is shortened, in the limit, the batch transfer scenario converges to one with instantaneous availability. By refining the subdivision of the working periods (maintaining the same time horizon, increasing the number of periods, therefore reducing their duration), one could have an item available for processing at the downstream station earlier. Thus, in the context of this study, it is assumed that instantaneous availability stands; work must satisfy the sequential order of activities, but are processed without being constrained by period boundaries. The investigation begins by establishing a relationship between the two-task flow-line and a slightly more tractable problem in Lemma 4.1.

Lemma 4.1 Solving the two tasks flow line scheduling problem with a general worker productivity model in the context of instantaneous availability is equivalent to maximizing the throughput from the least productive task (Hereafter the *MaxMin Problem*).

Proof in Appendix C.

In relation to the original flow line scheduling formulation, the MaxMin problem provides a suitable framework for characterizing switching points. Lemma 4.1 does not rely on a specific productivity model therefore it can be applied to more general worker performance models. In the next section, the focus is made on switching points for some productivity models.

4.2.1 MaxMin Formulation

The two-task *MaxMin* problem with two workers and no idle time is formulated as follows:

$$Maximize\ Q$$

s.t

$$Q \ \leq \ F_{i_1\, j_1}(t) + F_{i_2\ j_1}(T-t) \qquad (4-1)$$

$$Q \ \leq \ F_{i_2\ j_2}(t) + F_{i_1\, j_2}(T-t) \qquad (4-2)$$

$$0 \leq t \ \leq T \qquad (4-3)$$

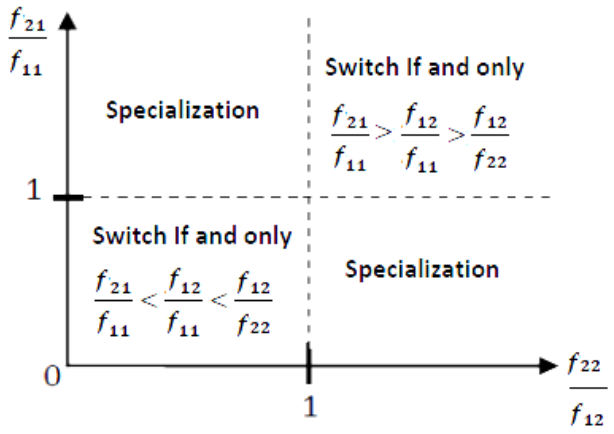
where $F_{i_p\, j_q}(t) = \int_0^t f_{i_p\, j_q}(u)du\ p \in \{1,2\}\ \ q \in \{1,2\}$ and $f_{i_p\, j_q}$ is the performance function (continuous function of time) that estimates the productivity rate of worker i_p on task j_q. Constraint (4-1) indicates that the throughput of the flow line cannot exceed what is produced on task j_1 while constraint (4-2) forces the same limitation with respect to task j_2. The equations only require that the integral of the productivity (learning) function to be defined over the time period considered. Hence, the approach can be used for steady, learning and/or forgetting, and stochastic productivity models. We note that we have implicitly assumed a specific initial assignment A_1 (worker i_1 at task j_1 and worker i_2 at task j_2) before any switching occurs. The other scenario with an initial assignment A_2 (worker i_1 at task j_2 and worker i_2 at task j_1) is similar in formulation and only requires a switch of the worker indices. Then, one can focus on one scenario (starting with assignment A_1 then switching to assignment A_2) since the outcomes for the other scenario (starting with assignment A_2 then switching to assignment A_1) can be easily deduced. The subsections that follow focus on defining optimal workflow policies and switching points.

4.2.2 Constant Production Rate

While not representative of the most general scheduling problems, the constant production rate assumption has been useful to establish production time standards. Given a constant production rate, lemma 4.1 leads to the result in Theorem 4.1.

Theorem 4.1 Within the two task flow line problem, with two workers when assuming non-null constant productivity rates, the decision to switch workers is determined by the ratios $\frac{f_{i_2 j_1}}{f_{i_1 j_1}}$, $\frac{f_{i_2 j_2}}{f_{i_1 j_2}}$ and the order of the ratios $\frac{f_{i_2 j_1}}{f_{i_1 j_1}}$, $\frac{f_{i_1 j_2}}{f_{i_1 j_1}}$, and $\frac{f_{i_1 j_2}}{f_{i_2 j_2}}$. The optimal the switching time is $t^* = \beta \times T$ where $\beta = \frac{f_{i_1 j_2} - f_{i_2 j_1}}{(f_{i_1 j_1} - f_{i_2 j_1}) + (f_{i_1 j_2} - f_{i_2 j_2})}$.

Proof in Appendix D.



(i_p and j_q is replaced by p and q for clarity).

Fig 4. 1. Two-task, two-worker flow-line optimal decision strategy with constant productivity

Theorem 4.1 is obtained by solving the MaxMin formulation. When assuming a steady productivity, the optimal policy between FAS and WSS depends on the workers' productivity rates. Figure 1 provides the optimal decision when maximizing throughput. For instance, with $\frac{f_{i_2 j_1}}{f_{i_1 j_1}} > 1$, $\frac{f_{i_2 j_2}}{f_{i_1 j_2}} > 1$ and $\frac{f_{i_2 j_1}}{f_{i_1 j_1}} > \frac{f_{i_1 j_2}}{f_{i_1 j_1}} > \frac{f_{i_1 j_2}}{f_{i_2 j_2}}$, the optimal policy is to adopt work sharing and switch workers. However, if the ratios do not satisfy that order, then specialization is optimal. The straightforward case with constant and common productivity rates on both tasks is given by Corrollary 4.1.

Corollary 4.1 For the two-task flow-line with two workers, when each worker has the same non-null constant production rate at both tasks it is always optimal to switch workers at *T/2* in order to maximize production.

Proof in Appendix E.

In the next subsection, more general models of performance are handled. An attention is specifically given to models that include learning since they better represent an important production system tradeoff.

4.2.3 Productivity with Learning

The existence and importance of learning effects on productivity has been known for some time (e.g., Conway and Schultz, 1959; Venezia, 1985; Cochran, 1960). In general, the theory of learning effects states that the time needed to produce a unit decreases with the processing of additional units. That is, with learning, worker performance improves. Theorem 4.2, below and corollaries 4.2-4.3 provide results for the two-task flow-line with learning.

Theorem 4.2 When maximizing throughput of a flow-line with two workers and two tasks considering a general productivity model that includes learning, the optimal switching time can be characterized as an inflection point on an integral (summation) of the productivity function. *Proof* in Appendix F.

The implications of Theorem 4.2 are that despite the dynamics of a flow-line with learning, the optimal switching point may be found in a straightforward manner using standard rules of calculus rather than requiring a full optimization.

Corollary 4.2 In a two-task flow line with two workers when using the hyperbolic learning model (Mazur and Hastie 1976) to represent productivity, the optimal switching time, when it exists, is a root of a polynomial function of degree four.

Proof in Appendix G.

Corollary 4.3 In the context of a productivity model with learning when maximizing the throughput in a two-task flow-line with two workers, if we assume that each worker is indifferent

to the two tasks and knowledge are not transferable, the optimal switching time of the workers, when needed, is at exactly half of the production time horizon. When knowledge can be transferred between tasks, it is always optimal to switch workers.

Proof in Appendix H

4.2.4 Stochastic Productivity

Stochastic effect has also been modeled as part of individual workers' productivity. One can model the processing time of a worker i_p at a station j_q as a random variable $\tau_{i_p\, j_q}$ that is the amount of work produced during a lapse t is $t/\tau_{i_p\, j_q}$.

Theorem 4.3 Within a flow-line with two tasks and two workers, when the processing time of worker i_p on task j_q follows an exponential distribution with parameter $\lambda_{i_p\, j_q}$ the decision to switch workers depends on the value of the ratios, $\frac{\lambda_{i_2\, j_1}}{\lambda_{i_1\, j_1}}$, $\frac{\lambda_{i_2\, j_2}}{\lambda_{i_1\, j_2}}$ and the order of the ratios, $\frac{\lambda_{i_2\, j_1}}{\lambda_{i_1\, j_1}}$, $\frac{\lambda_{i_1\, j_2}}{\lambda_{i_1\, j_1}}$, and $\frac{\lambda_{i_1\, j_2}}{\lambda_{i_2\, j_2}}$.

Proof in Appendix I.

The decision graph in Figure 3.1 holds when substituting $\lambda_{i_p\, j_q}$ for $f_{i_p\, j_q}$. The case of two tasks reveals the complexity of finding an optimal policy between FAS and WSS when the objective is to maximize overall production. For some instances of the distribution of workers' skills, the maximization of the throughput can be accomplished with an increase in flexibility (cross training) of the production system. In the next section, the case of a flow-line with three tasks is investigated.

4.2.5 Flow Lines with Three Tasks and Two Workers

The three task serial system is viewed as a system with two groups of tasks. By only allowing adjacent tasks to be grouped as a single task, two configurations are possible; configuration A (A_1 and A_2 being respectively the first and second group of tasks in configuration A) and configuration with group tasks B (B_1 and B_2 being respectively the first

and second group of tasks in configuration B). Since these configurations are different representations of the same flow line, the maximum throughput should be reached when adopting the optimal strategy in each configuration. It is assumed that:

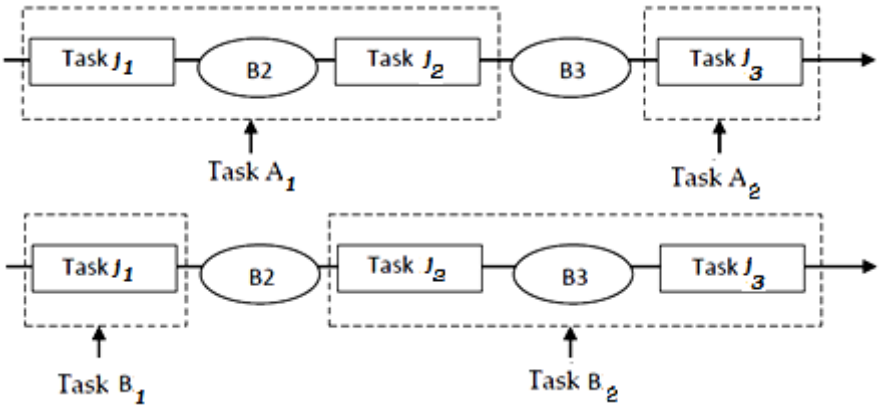


Fig 4. 2. Three-Task Flow-Line Configurations

- The worker assigned to task A_1 or task B_2 will adopt a pull system strategy (the start of a job at the first task of the group tasks is triggered by the completion of another at the second task of the same group tasks).
- During a production stint, a worker is assigned at most to a group of tasks. A group of tasks cannot be occupied by more than one worker.

While the worker index remains, i_p with $p \in \{1, 2\}$, the task index is j_q with now $q \in \{1, 2, 3\}$. Within configuration *A* and configuration *B* and under the two assumptions described above, the productivity of each worker for each group of tasks can be defined. Table 3.1 provides the critical values for switching in a manner similar to that in Figure 3.1.

Theorem 4.4 below will show that configurations *A* and *B* will yield the same optimal objective value. Figure 4.4 provides a numerical illustration of the optimal switching time for the two configurations when the optimal strategy is considered, and with workers productivities rate satisfying:

$$\begin{cases} y_{i_1 j_1} = 1, y_{i_2 j_1} = \alpha * y_{i_1 j_1}, T = 20 \text{ periods} \\ y_{i_1 j_2} = \alpha^2 * y_{i_1 j_1},\ y_{i_1 j_3} = \sqrt{\alpha} * y_{i_1 j_1} \\ y_{i_2 j_2} = \alpha^2 * y_{i_2 j_1},\ y_{i_2 j_3} = \sqrt{\alpha} * y_{i_2 j_1} \end{cases}$$

Table 4. 1. Critical Values for Three-Task Flow-Line Switching Conditions of figure 4.2.

	Two tasks with Exponential Processing Time	Three tasks - configuration A		Three tasks - configuration B	
		Constant Production Rate	Exponential Processing time	Constant Production Rate	Exponential Processing time
$f_{i_1\, j_1}$	$\lambda_{i_1\, j_1}$	$\frac{f_{i_1\, j_1} f_{i_1\, j_2}}{f_{i_1\, j_1} + f_{i_1\, j_2}}$	$\frac{\lambda_{i_1\, j_1} \lambda_{i_1\, j_2}}{\lambda_{i_1\, j_1} + \lambda_{i_1\, j_2}}$	$f_{i_1\, j_1}$	$\lambda_{i_1\, j_1}$
$f_{i_1\, j_2}$	$\lambda_{i_1\, j_2}$	$f_{i_1\, j_3}$	$\lambda_{i_1\, j_3}$	$\frac{f_{i_1\, j_2} f_{i_1\, j_3}}{f_{i_1\, j_2} + f_{i_1\, j_3}}$	$\frac{\lambda_{i_1\, j_2} \lambda_{i_1\, j_3}}{\lambda_{i_1\, j_2} + \lambda_{i_1\, j_3}}$
$f_{i_2\, j_1}$	$\lambda_{i_2\, j_1}$	$\frac{f_{i_2\, j_1} f_{i_2\, j_2}}{f_{i_2\, j_1} + f_{i_2\, j_2}}$	$\frac{\lambda_{i_2\, j_1} \lambda_{i_2\, j_2}}{\lambda_{i_2\, j_1} + \lambda_{i_2\, j_2}}$	$\lambda_{i_2\, j_1}$	$\lambda_{i_2\, j_1}$
$f_{i_2\, j_2}$	$\lambda_{i_2\, j_2}$	$f_{i_2\, j_3}$	$\lambda_{i_2\, j_3}$	$\frac{\lambda_{i_2\, j_2} \lambda_{i_2\, j_3}}{\lambda_{i_2\, j_2} + \lambda_{i_2\, j_3}}$	$\frac{\lambda_{i_2\, j_2} \lambda_{i_2\, j_3}}{\lambda_{i_2\, j_2} + \lambda_{i_2\, j_3}}$

When varying the value of parameter α the two configurations yield the same maximum throughput. However, as illustrated in Figure 4.4, it is important to note that the optimal switching *time* does not occur at the same point in these two configurations. Moreover, for some values of α ($1.4 \leq \alpha \leq 1.8$, a work-sharing policy with a switch of workers is optimal in one configuration (configuration B) while a specialization policy is optimal for the other configuration (configuration A).

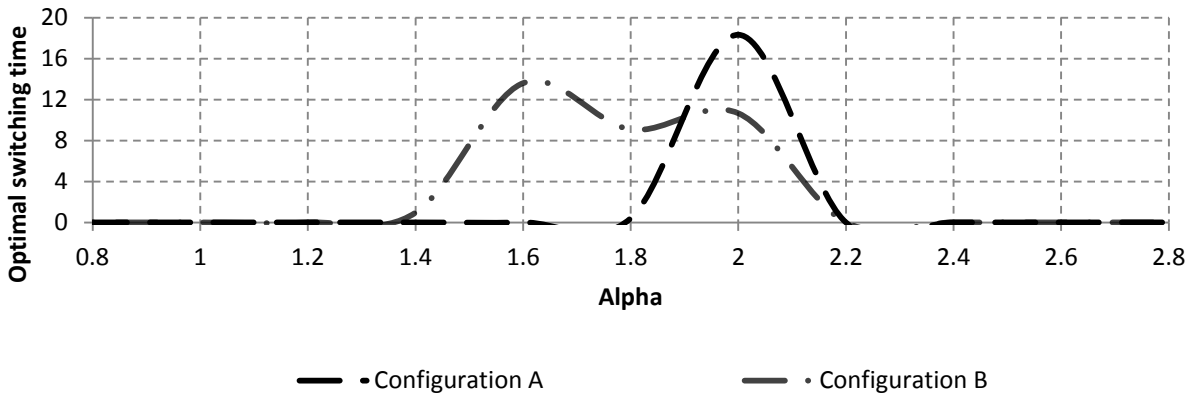


Fig 4. 3. Optimal Switching Time In Three-Task Flow-Line

The maximum throughput value, as the objective of the MaxMin problem, is established in section 4.3 with a general number of workers and tasks.

4.3 General *I*-Worker and *J*-task Case

Within the less complex configurations, the number of switches and optimal switching time could be characterized using the MaxMin problem analogy. Since this approach does not generalize directly, next is described a means of reducing the complexity of the general scheduling problem with I workers and J tasks during T periods of production. While considering a discrete number of periods in the time horizon, T represents common workplace practice, this directly is a major source of complexity in math program formulations. Indeed, schedules differentiate themselves by the assignments and their duration. Thus, knowledge about the maximum possible number switches has the potential to reduce the complexity of the corresponding math programming representation. The study of the two-task case illustrates this acutely, in that at most one change point is needed to obtain optimal output. In that case, there are at most two *stints* of production, although not necessary with the same duration. With this approach in mind, we consider the number of necessary change points for the general case of I workers, J tasks and time horizon T in Theorem 4.4.

Theorem 4.4 There exists an optimal schedule that requires no more than $J-1$ change points for the throughput maximization problem of the I workers and J tasks flow line. That is, the schedule is at most a sequence of $min(J,T)$ intervals (or stints) of assignments.

Proof in Appendix J.

Theorem 4.4 states that when developing schedules for I workers and J tasks, there is no loss of generality to only consider $min(J,T)$ assignments. That is, the duration of each interval can be part of the optimization problem, or can be given as input of the scheduling problem. Splitting the production time in T periods of production provides unnecessary change points, but it also affects the search performance of algorithms by decreasing the number of epochs at which a decision has to be made.

When examining the dynamics of the three-task case with two workers, it has been noticed that different configurations of the same flow line leads to the same output when the optimal strategy is considered. Theorem 4.5 formalizes this observation and provides the least-upper-bound (supremum) for the optimal output of the flow line scheduling problem by extending the general MaxMin problem analogy. In the MaxMin problem, the fundamental property of the flow-line is preserved. That is, the total production cannot exceed what has been produced at any one task in the flow-line. Tasks are in a sense, competing for workers, however, the objective is to maximize the production of the least productive task. The MaxMin problem for I workers and J tasks is formulated as follows:

Data

J Number of tasks $j = 1 \ldots 2$

I Number of workers $i = 1 \ldots 2$

T The production horizon

Θ_{ij} Set of parameters for worker i on task j

Matrix

O_{ijt} The output of worker i assigned to task j for t periods

$$O_{ijt} = \int_0^t f_{ij}(u, \Theta_{ij})du \qquad i = 1 \dots I \quad j = 1 \dots J \quad t = 0 \dots T$$

f_{ij} is a general model of productivity. Nembhard and Shafer (2007) use the same integral representation to consider the output through a given time period.

Decision Variables

m_{ijt} is a binary variable which equals 1 if worker *i* is assigned to task *j* for *t* periods.

Variables

F_{ij} is the output on task *j* from worker *i*

O_{ijt} is the production on task *j* from worker *i* when assigned for *t* periods.

The MIP is*:*

$$Maximize\ Q$$

s.t

$$Q \leq \sum_{i=1}^{I} F_{ij} \qquad i = 1 \dots I \qquad (Q-2)$$

$$F_{ij} = \sum_{t=0}^{T} m_{ijt} * O_{ijt} \qquad j = 1 \dots J \quad i = 1 \dots I \qquad (Q-3)$$

$$\sum_{j=1}^{J}\sum_{t=0}^{T} m_{ijt} * t = T \qquad i = 1 \dots I \qquad (Q-4)$$

$$\sum_{i=1}^{I}\sum_{t=0}^{T} m_{ijt} * t \leq T \qquad j = 1 \dots J \qquad (Q-5)$$

$$\sum_{t=0}^{T} m_{ijt} \leq 1 \qquad j = 1 \dots J \quad i = 1 \dots I \qquad (Q-6)$$

Constraint $(Q-2)$ is a standard surrogate for maximizing the least productive station. The output from worker *i* on task *j* is defined in constraint $(Q-3)$ for a given length of the assignment. Constraint $(Q-4)$ ensures that there is no idle time and constraint $(Q-5)$

stipulates that the total time a task is performed cannot exceed the time horizon *T*. Constraint $(Q-6)$ limits the worker's stint per task to a maximum of one.

Theorem 4.5 When solving the flow line assignment problem with *I* workers and *J* tasks, the optimal throughput can be obtained by solving the *MaxMin* formulation with the objective to maximize the least productive station.

Proof in Appendix K.

Theorem 4.5 stipulates that the throughput of the MaxMin problem, which can often be solved exactly, gives the optimal throughput of the flow line scheduling problem. It is important to notice that the corresponding schedule to such an exact MaxMin solution is not generally suitable as a solution to the original problem. However, the quality of any algorithm used to solve the original flow-line problem can be evaluated by comparing the objective value obtained to the MaxMin problem's throughput.

Conclusions

This study augments extant knowledge on optimal scheduling policies in serial systems with a general productivity model. A work sharing policy has been discussed to partially account for worker differences while a FAS policy is useful when the workforce is homogeneous. The impact of within worker variability on these of production systems has received relatively little previous attention. While there is a general understanding that WSS has advantages with a heterogeneous workforce, such as flexibility in the production system, it has been less clear whether individual worker variability with WSS will outperform FAS. An analogy between flow line scheduling and a MaxMin problem (maximization of the least productive task) is established. In the case of two or three tasks, the equivalent MaxMin problem provides a tractable framework to study the optimality of WSS and FAS within workers and between workers variability. Under some levels of worker performance FAS is the preferred policy. That is, a higher level of flexibility in some situation comes with a decrease in throughput. In a two or three task flow-line, the frontier between WSS and FAS optimally as a function of workers' productivity parameters is defined. Switching time and the minimal number of switches when a WSS policy is optimal are discussed

The current study intends to address broad managerial questions such as whether or not to cross train the workforce, and when cross training, how often should workers switch tasks, generally. This direction is distinct from other approaches that for instance view systems dynamically, such as the work from Bartholdi et al. (1999), Bartholdi and Eisentein (1996), and Bartholdi et al. (2001) on the use of *bucket brigades* in production systems. Also, Armbruster and al. (2007) propose an extension of the bucket brigades with worker learning, however results address post-transient behavior. The current study directly addresses managerial worker-scheduling decisions during the transient phase when learning has a significant impact on individual and system performance. Results show that implementing the optimal switching time will lead to optimal throughput. And, despite the inherent complexity of including learning in serial production, the optimal switching time can be determined using a calculus rule.

The study and the methodology presented here can be conceivably applied using any practical productivity model. The results suggest that in a flow line production system, depending on the distribution of the workforce skills, a cross training policy could be implemented without decreasing throughput. There are a wide range of suitable rationales for cross training a workforce in production systems generally. Nonetheless, unlike parallel systems where cross training comes along with some costs with respect to quantitative system output, flow-line systems exhibit cases where implementing cross training policy can lead to a maximization of throughput. When worker switching is implemented in a flow line, it is possible to achieve both higher flexibility and higher production output. This may provide production managers additional rationale for implementing cross training, especially when addressing the assignment and scheduling of workers to tasks. Optimal switching times for 2-3 tasks and two workers are illustrated. Extending this approach to larger potential problem instances is of considerable interest for future research.

CHAPTER 5

WORKFORCE HIRING AND CROSS-TRAINING IN PARALLEL AND SERIAL SYSTEM WITH HUMAN PERFORMANCE VARIABILITY

5.1 Introduction

Chapter 3 and Chapter 4 were devoted to the study of parallel and serial system with the aim of providing insights on the structure of the optimal policy considering worker performance variability. These results were established with the only constraint being to maximize throughput. This chapter aims to provide more insight on the policy to apply when flexibility and cross-training level are required. Five policies that rank, select and use a learning and forgetting-based assignment are implemented. The impact of cross training magnitude, level of multi-functionality and workforce heterogeneity on the policies performance is of interest. The chapter is organized as follows. Section 5.2 is devoted to a presentation of the methodology. Section 5.3 focuses on the selection framework describing the implemented policies and the two considered production scenarios. Section 5.4 summarizes the results of the simulation design and discussion of the outcomes is presented in section 5.5.

5.2 Methodology

In order to examine the question of how to select and assign the workforce at a production system, several simulations are conducted considering five specific worker selection policies described below. The decision-making can occur in distinct stages, namely worker selection, worker assignment, and worker scheduling, with selection involving the who question, assignment involving the question of what tasks, and scheduling involving the timing of when tasks are to be done. The considered policies involve the first two of these stages. In the first stage, workers are selected using a specific policy. In the second stage a math programming formulation is solved to optimize worker assignments. The performance of a worker on a task is

estimated using a productivity model that includes L/F characteristics. The different ranking methods used rely either partially or completely on the L/F model parameters. The math program that constitutes the second stage of the policy is formulated assuming knowledge of the L/F model's parameters. By examining the five policies, the aim is to determine the degree to which they improve productivity.

In addition to the specific policies considered, the effects of the ratio of generalists to specialists is studied, where a specialist describes a worker who is assigned to a single task during the entire production horizon, while a generalist is a cross trained worker that must performed a specified number of tasks. The number of tasks to be executed by all generalist workers is determined by the level of multi-functionality, which is only applied to generalist workers. When the level of multi-functionality equals the number of tasks, we describe this scenario as having full flexibility. An associated experimental factor is to examine the impact of the level of workforce heterogeneity on productivity. The research objective is to identify preferred operating levels for these factors in a setting having independent and dependent tasks.

As in Table 5.1, the simulation considers 5 policies × 5 worker pools × 3 levels of the ratio specialist to generalist × 2 levels of multifunctionality × 2 levels of workforce heterogeneity, which results in 300 experimental treatments. For each treatment, 10 replications are run for 4 months of simulated production time.

Parallel and serial system scenarios are considered. These systems are each of interest but not directly comparable to one another so they constitute two separate cases. In parallel system four independent tasks are considered, hence work on a specific task may be completed independently of the other three tasks. The serial system also consists of four tasks but jobs must progress through each of the four stations in turn in order to complete units of production.

Table 5. 1. Simulation Design Factors and Levels

Factors	*Levels*
Selection Policy	Random (policy *A*) MaxiMax prior expertise (policy *B*) MaxiMin prior expertise (policy *C*) MaxiMax Expected Output (policy *D*) MaxiMin Expected Output (policy *E*)
Specialist/Generalist Ratio (SGR)	0/4, 1/3, 2/2
Generalist Multi-functionality (GM)	1, 2
Workforce Heterogeneity (WH)	1, 3
Pool of Workers	5 random pools of workers

Table 5. 2. L/F parameter means, variance-covariance matrix

$$\mu = \begin{bmatrix} 1.448 \\ 1.963 \\ 2.0446 \\ 2.269 \end{bmatrix} \begin{matrix} \log k \\ \log p \\ \log r \\ \alpha \end{matrix} \qquad \Sigma = \begin{matrix} \log k \\ \log p \\ \log r \\ \alpha \end{matrix} \begin{bmatrix} 0.0045 & 0.0167 & 0.0191 & -0.0214 \\ 0.0167 & 0.4621 & 0.2443 & -0.2380 \\ 0.0191 & 0.2443 & 0.1905 & -0.2326 \\ -0.0214 & -0.2380 & -0.2326 & 0.3638 \end{bmatrix}$$

$$\qquad\qquad\qquad\qquad \log k \qquad \log p \qquad \log r \qquad \alpha$$

5.3 Worker Selection Framework

The simulation begins with a pool of n workers based on empirical data from the literature (Nembhard and Osothsilp, 2005; Shafer *et al*, 2001). The data span a 6-month period from the final inspection stations of a car radio assembly process. A statistical summary of these data are given in Table 5.2. From this general pool we will select workers based on each of the five selection policies, forming sub-pools of four workers. Depending on the simulation treatment, workers may be split into two groups: a group of specialists and a group of generalists. If the specialist to generalist ratio is x/y with $x + y = 4$, the group of specialists is size x and the generalists, y. Once the workers are selected the remaining $n - 4$ workers are assumed to perform some tasks outside of the scope of this work setting or are for example absent, sick or on vacation.

Workers' L/F parameters values

In this study workers' L/F characteristics are represented by the *hyperbolic-recency* model of L/F described in Appendix A. The four parameters that comprise the model are the steady-state productivity rate, the learning rate, the forgetting rate, and prior expertise. The approach considered enables managers to estimate workers' L/F parameters using production data from past or previous work experience. For each worker, the four L/F parameters are estimated by fitting the model to empirical production data. This approach can be applied to a wide variety of L/F models and similarly the estimation process may be conducted via any appropriate methodology. (Yelle, 1979) and (Dar-El, 2000) provide several alternatives for the parameter estimation as well as model choice. We minimize estimation error by choosing a productivity model that tends to lead to smaller errors (see e.g., Nembhard and Osothsilp, 2001).

Data from 26 workers were fit to the hyperbolic-recency model to estimate individual L/F parameters. In our simulation, a larger population of 100 workers (that is $n = 100$) was generated using the multivariate distribution of parameters derived from the 26 workers. For the 100 workers, the set of $k, p, r,$ and α parameters is used as population parameters and is generated from the multivariate distribution of $\log k, \log p, \log r,$ and $\alpha,$ with mean vector and covariance matrix shown in Table 5.2.

The five factors are outlined in Table 5.1. The first factor is the *policy* used to select workers. The broad question is whether recruitment should prioritize homogeneous capabilities or single task efficiency and how these policies perform with a given level of flexibility among the workers (i.e., the number of generalists and multi-functionality level). The five policies are enumerated:

The effective baseline policy is to randomly select workers from the larger pool of *n* workers. This subset will then be assigned to tasks based optimizing the assignments of the workers selected. It is noticed that with respect to productivity, it is most common to select subsets of workers within an organization without knowledge of L/F characteristics. Thus, this random policy is a straightforward way to simulate outcomes based on common practices.

For the *MaxiMax–prior expertise* policy, workers are ranked based on the level of prior expertise, parameter p in the hyperbolic-recency model (see Appendix A). The highest rank is

associated with the worker with the highest value of *p* from any of the four tasks, the next highest is for the second highest value of *p* at any of the four tasks, and so on.

For the *MaxiMin–prior expertise* policy, workers are also ranked based on the level of prior expertise, however the hierarchy depends on the minimum level of initial expertise. That is, the highest rank is associated with the worker that has the greatest minimal value of *p* among the four tasks.

In the *MaxiMax Expected Output* policy, workers are ranked based on an overall estimate of the output they can be expected to produce at the four tasks, with the highest rank given to the worker with the greatest projected output across the four tasks. We maximize across each of the workers' best tasks. The estimates are based on applying all four parameters of the L/F model over the time horizon.

In the *MaxiMin Expected Output* policy, for each worker we first determine the minimum expected output across the four tasks. The workers are then ranked by maximizing these task-performance minima. That is, we maximize across each worker's least productive tasks.

One can posit that the two *MaxiMax* policies may favor workers that are more efficient, and are in some sense greedy approaches. That is, while they are highly efficient at one task, they may perform poorly on others. The *MaxiMin* policies attempt to account for this possibility by considering each worker's worst task.

The second factor, after policy, is the *specialist to generalist ratio,* which is defined as the ratio of the number of specialist workers to the number of generalist workers in the selection. A specialist will perform a single task during the entire production horizon. A generalist worker will share production time among several tasks depending on the specific level of multi-functionality. It is assumed that there will be a worker on each task at all times and once a worker leaves a task, a substitute worker from the sub-pool of selected workers is moved to that task instantaneously. It is anticipated that there should be a threshold for the specialist to generalist ratio. The number of generalist workers can be viewed as a form of flexibility in the system. Full flexibility has been widely discussed in the literature (e.g., Hopp et al. 2004) to often lead to sub-optimal production. Three levels of the specialist to generalist ratio 0/4, 1/3 and 2/2 are considered. For example, a ratio 1/3 represents that among the four workers selected one

worker will be a specialist performing one task, and the three others will be generalists performing multiple tasks.

The third factor, *multi-functionality,* is defined as the number of tasks a generalist worker will perform. The specialist to generalist ratio and multi-functionality together indicate the degree of flexibility in the production system with the latter affecting only the generalist workers. While higher levels of multi-functionality may lessen productivity, its impact on serial systems is less clear. Levels of 1 and 2 are selected, which when combined with the specialist to generalist ratio can form a wide range of scenarios. For example, with a ratio of 1/3, the level of multi-functionality will be at most three since a specialist is said to perform only one task and must not be idle.

The fourth factor, the *workforce heterogeneity,* is represented by the variance-covariance matrix of the four hyperbolic–recency model parameters. The variance-covariance matrix is $\delta \times \Sigma$, where Σ is given in Table 4.2. Worker characteristics are sampled from the multivariate distribution by considering $\delta = 1$ and $\delta = 3$. A workforce heterogeneity, $\delta = 1$, corresponds to a case when we maintain the empirically observed level of variability. By increasing heterogeneity to 3, one can examine whether or not heterogeneity has an impact on the effectiveness of the policies considered.

The fifth factor is the *pool of workers,* which contains all workers to be selected including both generalists and specialists. This is a random factor to allow for multiple observations of the other fixed factors. Since there is significant variation among workers we randomly sample five worker pools, each one with a population of 100 workers.

LFBA model under two production scenarios

The L/F-based assignment (LFBA) defines the second stage after selecting the workers. Depending on the production system considered, independent or dependent tasks, a mixed integer programming model is solved to determine the optimal assignment of the four selected workers. Models 1-2 determine the worker assignment for the parallel and serial systems, respectively.

Model 1: Parallel systems

j task index $j = 1,2, \ldots, 4$

i worker index $i = 1,2, \ldots, 4$

t index of production periods $t = 1,2, \ldots, T$

M A large positive number

P_{ijtp} The output of worker i assigned to task j at time t with a previous cumulate experience of p periods.

$$P_{ijtp} = \int_{p}^{p+1} f_{ij}(u, \Theta_{ij}) du \quad for\ p < t \quad i = 1 \ldots I \quad j = 1 \ldots J \quad t = 1 \ldots T$$

$$P_{ijtp} = 0 \quad for\ l \geq t \quad i = 1 \ldots I \quad j = 1 \ldots J \quad t = 1 \ldots T$$

f_{ij} a general model of productivity.

Θ_{ij} set of parameters of worker i on task j.

Binary decision Variables

z_{ijtp} equals 1 if worker i with p periods of experience on task j is assigned to that same task at time t and 0 otherwise.

x_{ijt} equals 1 if worker i is assigned to task j at time t and 0 otherwise.

$wgen_i$ equals 1 if worker i is selected as a generalist and 0 otherwise.

$wspe_i$ equals 1 if worker i is selected as a specialist and 0 otherwise.

Binary Variables

$ygen_{ij}$ equals 1 if worker i considered as a generalist performs task j and 0 otherwise.

$yspe_{ij}$ equals 1 if worker i considered as a specialist performs task j and 0 otherwise.

Non negative Variables

S_{jt} the output from task j at period t.

B_{jt} the level of buffer j at period t.

P_{ijtp} the predict output of worker i on task j at time t with cumulative assignment p according to the model of productivity used.

O_{ijt} the output of worker i on task j at time t.

$$Maximize \sum_{j}^{4} \sum_{t}^{T} S_{jt}$$

$$S_{jt} = \sum_{i}^{4} O_{ijt} \qquad j = 1..4 \quad t = 1..T \qquad (5-1)$$

$$O_{ijt} = \sum_{p \le t} z_{ijtp} * P_{ijtp} \qquad i = 1..4 \quad j = 1..4 \quad t = 1..T \qquad (5-2)$$

$$\sum_{p \le t} z_{ijtp} \le x_{ijt} \qquad i = 1..4 \quad j = 1..4 \quad t = 1..T \qquad (5-3)$$

$$\sum_{p \le t} z_{ijtp} \times p \le \sum_{q \le t} x_{ijq} \qquad i = 1..4 \quad j = 1..4 \quad t = 1..T \qquad (5-4)$$

$$\sum_{i=1}^{4} x_{ijt} = 1 \qquad j = 1..4 \quad t = 1..T \qquad (5-5)$$

$$\sum_{j=1}^{4} x_{ijt} = 1 \qquad i = 1..4 \quad t = 1..T \qquad (5-6)$$

$$\sum_{t} x_{ijt} = T \times yspe_{ij} + \frac{T}{M} \times ygen_{ij} \quad i = 1 \dots 4 \quad j = 1 \dots 4 \qquad (5-7)$$

$$yspe_{ij} + ygen_{ij} \ge x_{ijt} \quad i = 1 \dots 4 \quad j = 1 \dots 4 \qquad (5-8)$$

$$\sum_{j} ygen_{ij} = M \times wgen_{i} \qquad i = 1..I \qquad (5-9)$$

$$\sum_{j} yspe_{ij} = wspe_{i} \qquad i = 1..I \quad j = 1..J \qquad (5-10)$$

$$yspe_{ij} + ygen_{ij} \le 1 \quad i = 1 \dots 4 \quad j = 1 \dots 4 \qquad (5-11)$$

$$wspe_{i} + wgen_{i} = 1 \quad i = 1 \dots 4 \qquad (5-12)$$

$$\sum_i wgen_i = 4 - N \qquad (5-13)$$

$$\sum_i wgen_i + \sum_i wspe_i = 4 \qquad (5-14)$$

Within the parallel production scenario, the objective is to maximize the output from all four tasks. Constraint $(5-1)$ to constraint $(5-4)$ define the output at each of the four tasks, the output from all four workers selected, and the learning as the result of cumulative assignments. Constraints $(5-5)$ and constraint $(5-6)$ ensure that at each time *t*, a task is occupied by exactly one worker and a worker is assigned to exactly task. Constraint $(5-7)$ and constraint $(5-8)$ define the time spent at each tasks depending on the worker type; specialist or generalist. Constraint $(5-9)$ and constraint $(5-10)$ establish the link between the variables $ygen_{ij}, yspe_{ij}$ and $wgen_i, wspe_i$ defining the worker type; specialist or generalist. Constraint $(5-11)$ to constraint $(5-14)$ ensure that a task is either performed by a specialist or a generalist, a worker is either a specialist or a generalist, and the selected workers are divided into two groups.

Model 2: Serial systems

This second scenario is devoted to the case of the flow line with four tasks sequenced one after another. A unit must progress through the four tasks sequentially, and this sequence is constant for all jobs. Thus, the objective is to maximize the number of products that are completed at the final task-station. In addition to sequencing, serial systems generally use buffers that constrain the performance of the worker assigned to a downstream station. Thus, Model 1 can be applied with the inclusion of Equation $(5-15)$, with buffers $B_{j\,t}$.

$$B_{j\,t} = B_{j\,t-1} + \sum_i^4 O_{ij-1\,t} - \sum_i^4 O_{ijt} \qquad j = 2 \ldots 4 \quad t = 2 \ldots T \qquad (5-15)$$

The objective function can be replaced by summing only the output from the last station, as follows.

$$Maximize \sum_{t}^{T} S_{4t}$$

Analysis methodology

The significance of the experimental factors is determined by fitting a mixed-effects model (see e.g., Pinheiro and Bates 2000). Fixed and random effects are represent in the form $y = X\beta + Z\gamma + \varepsilon$, where y denotes the vector of observed values, X is the known fixed effects design matrix, and β is the unknown fixed effects parameter vector. $Z\gamma$ is the random component of the mixed model with Z being the known random effects design matrix and γ the vector of unknown random-effects parameters. As X contains the variables constructed for fixed effects, Z contains the indicator variables for the fixed effects. The unobserved vector of independent and identically distributed Gaussian random errors is represented by ε.

5.4 Results

The ANOVA of fixed effects for the parallel and serial systems output is displayed in Table 5.3. Table 5.4 provides 95% confidence intervals for output obtained at each treatment (i.e. for a given policy, specialist to generalist ratio, generalist multi-functionality, and workforce heterogeneity level). Table 5.5 indicates the improvement in production obtained from the policies relative to the baseline random policy for both production systems. Results reported as significant in this section are at a 5% level.

Table 5.3 indicates that the factors policy, specialist to generalist ratio, and the workforce heterogeneity are significant in both production scenarios. The generalist multi-functionality is only significant in the parallel case. Interaction effects between policy and specialist to generalist ratio, and between policy and the workforce heterogeneity are also significant in both systems. The interaction between specialist to generalist ratio and workforce heterogeneity is only significant for the parallel system. The interaction effect between generalist multifunctionality and workforce heterogeneity is not significant in either scenario.

Figure 5.1.(left) and figure 5.1.(right) illustrate that the *MaxiMax P* policy outperforms the other policies in both parallel and serial systems. While *MaxiMin P* is significantly better than

MaxiMin O in the parallel tasks setting, the two policies do not differ significantly in the serial system. *MaxiMax O* did not outperform the baseline policy in either production system (Table 5.5). However, each of the other policies did outperform the baseline. The minimum percentage improvement for the *MaxiMin P* policy was 21.1 % in the parallel system and 12.1 % in the serial system.

Table 5. 3. Mixed effects ANOVA for simulation treatments

		Parallel output		Serial output	
Source	d.f	$F-value$	$Pr > F$	$F-value$	$Pr > F$
Policy	4	1006.39	0.0000*	1103.65	0.0000*
Spec. /Gen. Ratio	2	34.48	0.0001*	46.87	0.0000*
Gen. Multifunctionality	1	4388.32	0.0000*	6.84	0.0591
Workforce Heterog.	1	9794.24	0.0000*	2414.52	0.0000*
Policy* Spec. /Gen. Ratio	8	16.31	0.0000*	8.91	0.0000*
Policy* Gen. Multifunct.	4	25.88	0.0000*	16.07	0.0000*
Policy* Workforce Heterog.	4	379.22	0.0000*	116.89	0.0000*
Spec. /Gen. Ratio * Workforce Heterog.	2	5.84	0.0274*	4.01	0.0621
Gen. Multifunct.* Workforce Heterog.	1	0.66	0.4616	6.11	0.0687

*Significance at 5% level,

Gen. = Generalist, Spec. = Specialist, Hetero. = Heterogeneity, Multifunct. = Multi-functionality.

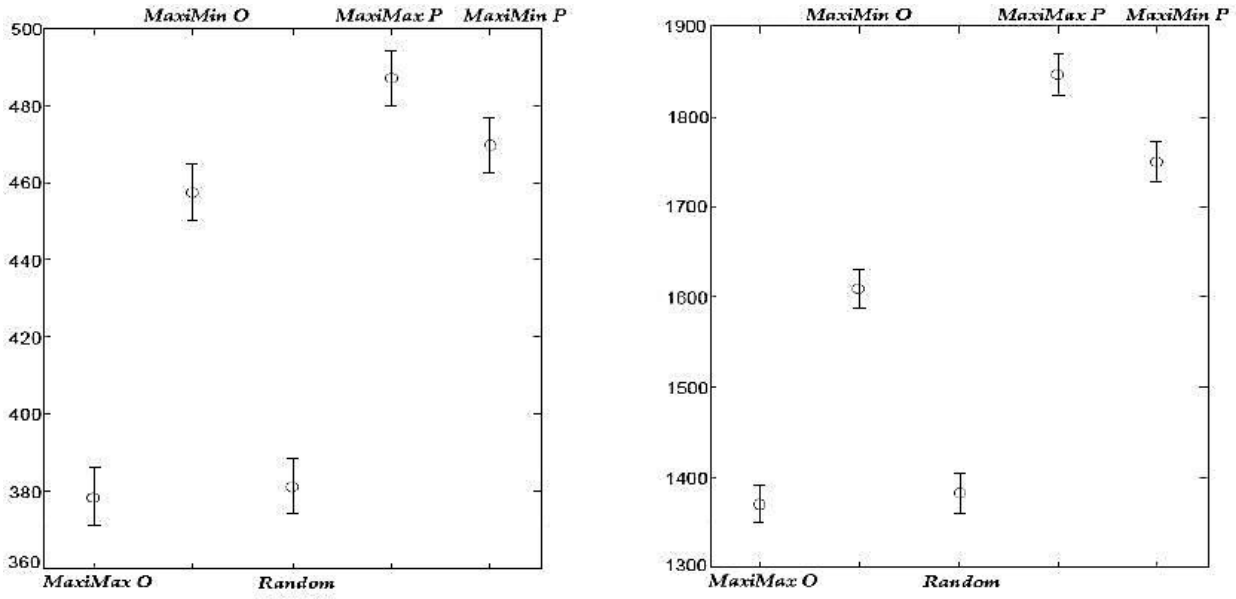


Fig 5. 1. Policy Main effects Serial (left) and Parallel (right) system

Figure 5.2.(right) shows that on average for the parallel task scenario, decreasing the specialist to generalist ratio leads to a decrease in production by 1% (from ratio 2/2 to ratio 1/3) and 2% (from ratio 1/3 to ratio 0/4).

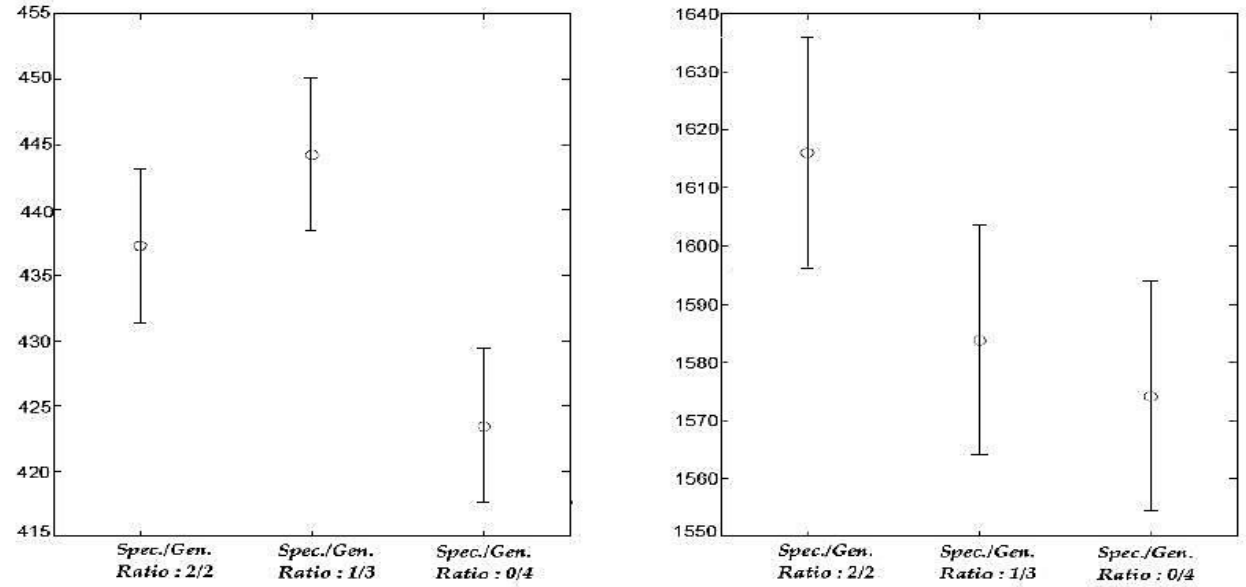


Fig 5. 2. Spec./Gen. Main effects Serial (left) and Parallel (right) system

The 2/2 specialist to generalist ratio performed significantly better than the 0/4 ratio. That is, on average, with a parallel configuration, maintaining some specialists may be preferred in systems that require cross training flexibility. That is, if flexibility is needed, it may be not only sufficient, but also desirable to only cross train a subset of the workers in the production process.

Table 5. 4. Percentage Improvement in Output Relative to Random Policy

		Parallel System				Serial System			
		Generalist Multifunctionality				Generalist Multifunctionality			
		1		2		1		2	
		Variance Heterogeneity				Variance Heterogeneity			
Spec./Gen	*Policy*	1	3	1	3	1	3	1	3
0/4	*MaxiMax P*	21.7	43.3	29.8	54.6	23.3	40.0	32.4	48.1
	MaxiMin P	13.7	27.2	23.8	41.6	14.3	24.4	26.8	37.3
	MaxiMax O	-0.6	1.8	-1.3	-3.2	-2.6	6.7	-0.9	-0.2
	MaxiMin O	8.9	9.1	12.8	28.8	10.3	8.8	18.5	38.1
1/3	*MaxiMax P*	21.7	39.5	21.1	42.1	18.7	39.4	12.1	31.2
	MaxiMin P	16.6	30.0	19.1	37.8	16.0	26.9	17.8	34.9
	MaxiMax O	-2.6	1.0	1.3	-5.6	-3.9	-0.6	4.5	0.5
	MaxiMin O	10.9	11.1	16.0	29.0	14.0	21.1	16.1	35.5
2/2	*MaxiMax P*	24.2	35.8	27.8	39.1	16.1	21.8	17.3	30.9
	MaxiMin P	16.6	30.0	23.2	38.3	13.8	23.3	19.3	22.6
	MaxiMax O	-1.6	-3.2	4.9	-1.1	-2.2	-2.3	-1.9	-5.1
	MaxiMin O	6.6	21.0	17.1	25.7	10.9	23.7	14.5	26.8

Figure 5.3. (left) and Figure 5.3 (right) illustrate that increasing the multi-functionality of generalist workers on average leads to a decrease in production. However, the decrease is only significant in the parallel task configuration. Figure 5.4 (left) and Figure 5.4 (right) show that higher levels of workforce heterogeneity leads to a 20% improvement in the serial system output and 17% in the parallel configuration.

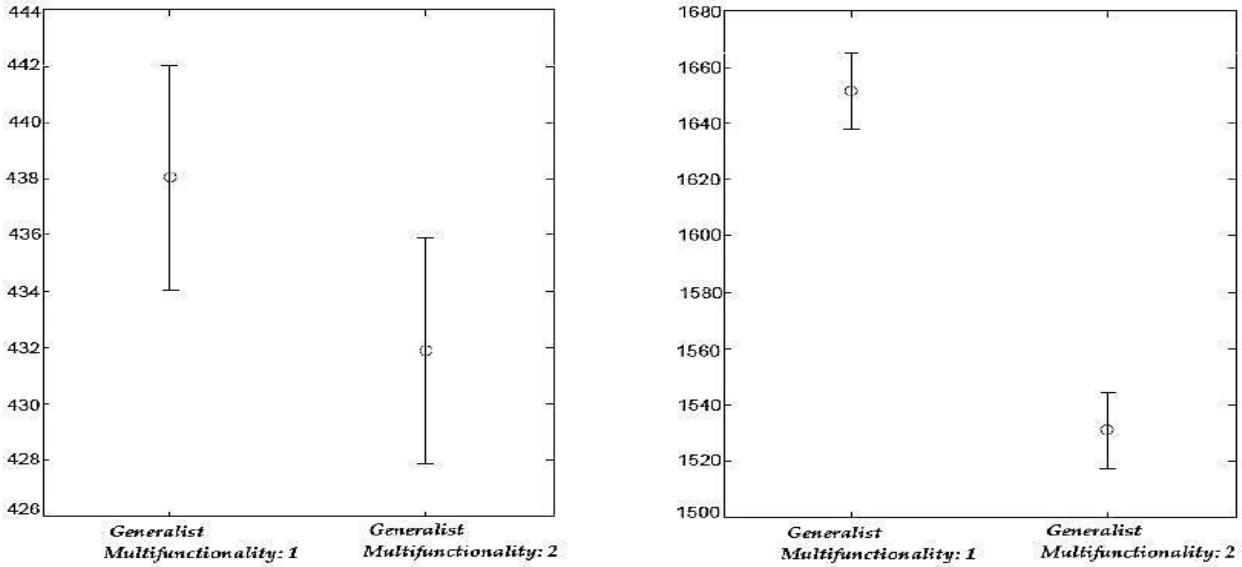


Fig 5. 3. Gen. Multifunct. Main effects Serial (left) and Parallel (right)

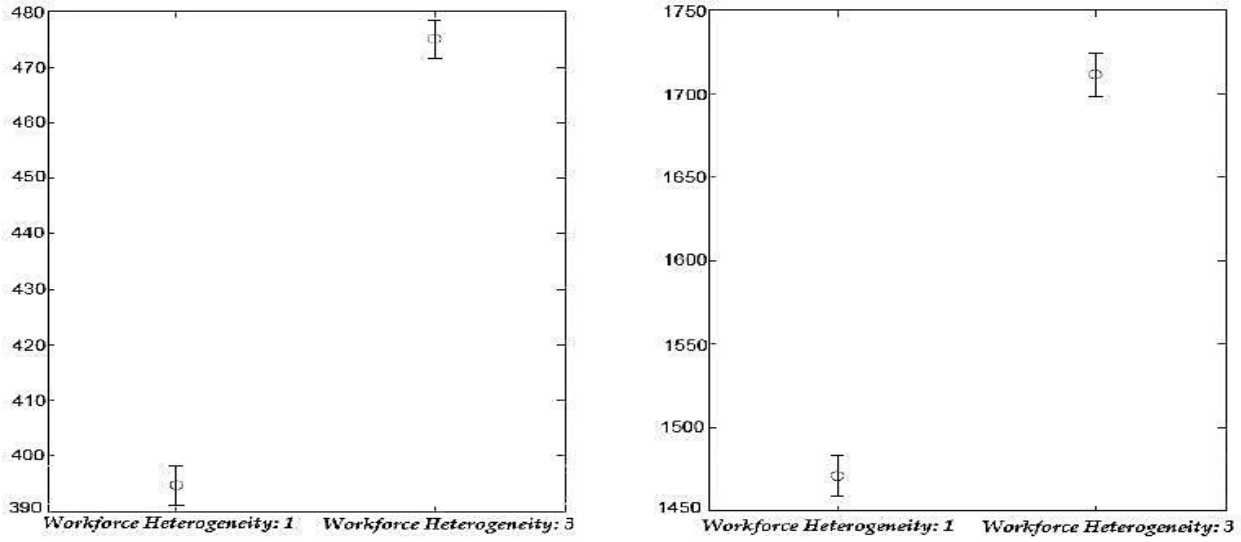


Fig 5. 4. Workforce Heterog. Main effects Serial (left) and Parallel (right) system

Table 5.3 indicates that the interaction between the policy factor and the other fixed factors, namely Specialist to Generalist Ratio, Generalist Multi-functionlaity, Workforce Heterogeneity, are significant in both configurations. Table 5.4 details these interactions for each treatment.

Figure 5.5 expands the results from Figure 5.1. That is, *MaxiMax P* outperforms the other policies followed by *MaxiMin P* policy and *MaxiMin O*, with *MaxiMax O* performing roughly similar to the baseline policy. Figure 5.5 (left) and figure 5.5 (right) illustrate that these policies have similar behavior. In parallel systems their performance decreases with a decrease in the specialist to generalist ratio. In the serial case *MaxiMax P* has higher performance with a decrease in the specialist to generalist ratio while there was no clear trend in parallel system.

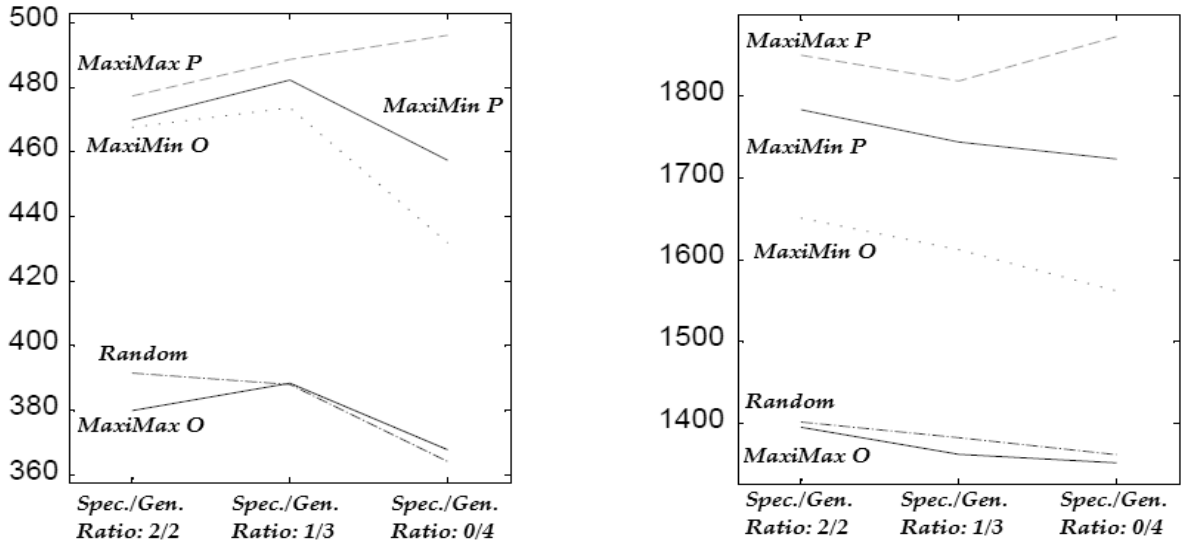


Fig 5. 5. Interaction plot Policy vs Spec./Gen. Ratio Serial (left) and Parallel (right) system

The interaction between policy and generalist multi-functionality is illustrated in Figures 5.6 (left) and figure 5.6 (right). For the parallel tasks, the performance of all policies decreases with an increase in generalist multi-functionality. For serial tasks the *MaxiMax P* and *MaxiMax O*, policy outputs decrease with an increase of generalist multi-functionality while *MaxiMin P* and *MaxiMin O* policies have the opposite evolution.

The interaction between the policy and workforce heterogeneity is shown in Figures 5.7 (left) and 5.7 (right) In both parallel and serial scenarios, an increase in workforce heterogeneity leads to improvement in each policy. This is consistent with results from Nembhard and Shafer (2008)

who showed that heterogeneity is often beneficial to overall output. However, the current results extend this notion to the performance of worker selection policies. Further figures 5.7 (left) and figure 5.7 (right) also suggest a multiplier effect wherein, the highest benefits from heterogeneity may also be associated with the better performing policies.

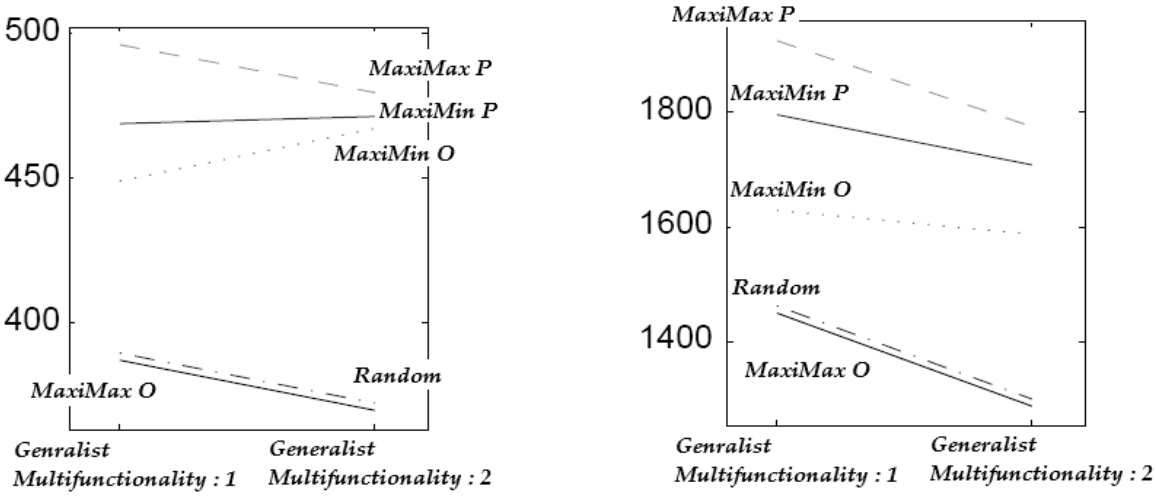


Fig 5. 6. Interaction Plot Policy vs Gen. Multifunct. Serial (left) and Parallel (right) system

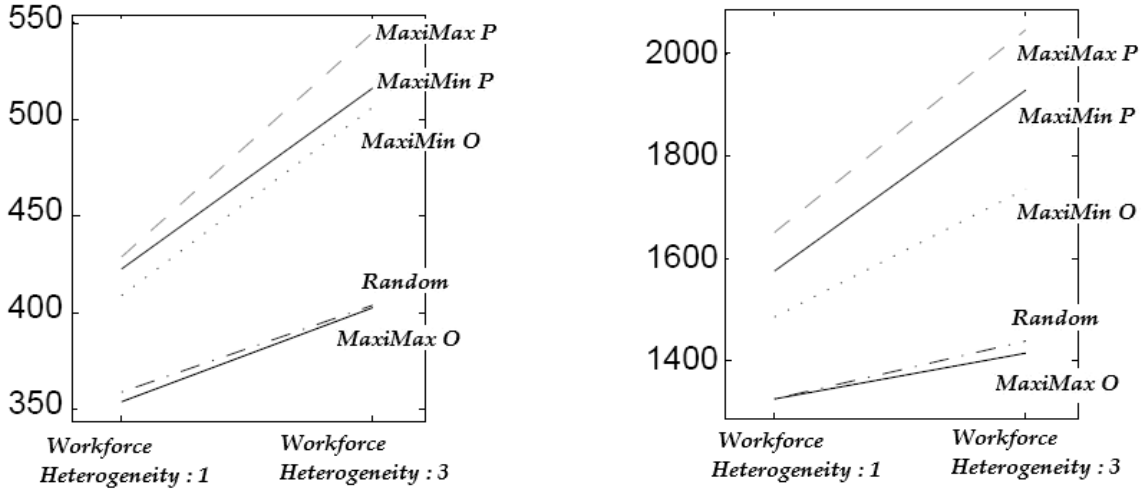


Fig 5. 7. Interaction Plot Policy vs Workforce Heterogeneity Serial (left) and Parallel (right) system

Table 5. 5. Average Improvement Of Policies Relative To Baseline

Policy	Parallel	Serial
MaxiMax P	34.5 %	28.9 %
MaxiMin P	27.2 %	23.6 %
MaxiMax O	-1.8 %	-1.3 %
MaxiMin O	16.3 %	21.0%

5.5 Discussion

Four policies that have the potentially to improve the process of selecting workers from a larger pool based on some limited available knowledge of worker capabilities are considered. With *MaxiMax O* and *MaxiMin O,* workers are ranked based on the knowledge of all four L/F parameters, and estimates of the expected output from each task. *MaxiMax P* and *MaxiMin P* rely solely on the single prior expertise parameter, *p*. Similar to results from Nembhard and Osothsilp (2005), these partially informed policies perform no worse than the baseline, and in some cases significantly outperform the baseline. Secondly, while Nembhard and Osothsilp suggest that basing worker selection and assignment on L/F characteristics can provide important improvements, it is further demonstrate that several straightforward policies that do not require instance specific optimization can produce such improvements. In this case, three of the four policies examined outperformed the baseline policy. Among this set, the *MaxiMax P* policy is perhaps the best performing of the four with respect to system throughput in both parallel and serial configurations.

Table 5.6 averages the relative improvements of each policy over the baseline. *MaxiMax P* and *MaxiMin P* policy results suggest that having significant prior expertise provides an effective head start on performance and learning. Equation (5-1) illustrates this notion, wherein *MaxiMax P* and *MaxiMin P* favor higher levels of prior expertise (a higher value of prior expertise with less variance among tasks in the case of the *MaxiMin P* policy).

$$\frac{dy_x}{dp} = \frac{kr}{(xR^{\alpha}{}_x + p + r)^2} \qquad (5-1)$$

That the MaxiMax *O* policy was not significant, while the other related policies were significant is at first somewhat surprising. However, it is acknowledged that the greedy policy selects workers based on the maximum output that can be obtained if the worker performed a single task during the entire simulation time. Selected workers will be most efficient at that task, but potentially inefficient at others. Since the goal is not to form teams of specialists, but rather groups with some flexibility, there are enough workers underperforming at their alternate tasks to offset gains made from their primary tasks. This illustrates that selection and assignment might ultimately be combined as a single process. However, the *MaxiMin O* policy effectively overcomes the limitations of the *MaxiMax O* policy by favoring workers with more homogeneous performance at the four tasks.

Effects of Specialist to Generalist Ratio

The specialist to generalist ratio allows investigating how both the policies and production systems perform with an increase in flexibility. Increasing flexibility will often involve increasing the number of generalists and decreasing the number of specialists. One can observed that higher levels of flexibility (i.e., more generalists) tend to reduce productivity, with a stronger effect in parallel systems. This is in concordance with analytical findings in Nembhard and Bentefouet (2012) that report theoretical optimality for specialization in parallel systems. The current result also illustrates that the productivity differences are roughly 5%, a small but potentially important difference in competitive industries. In the meantime, increase in flexibility can also positively affect system performance, however many studies have shown diminishing returns on labor flexibility (e.g., Park and Bobrowski, 1898; Malhortra et al., 1993; Fry et al., 1995; Campbell, 1999; Molleman and Slomp, 1999). In both parallel and serial systems, the lowest performance was obtained with full flexibility (i.e., four generalists). The generalists must rotate though all tasks, with the concomitant losses due to both learning and forgetting.

Effects of Generalist Multifunctionality

Higher levels of multifunctionality tended to reduce productivity in both production parallel and serial scenarios. However, the effect was only significant in the parallel case with a roughly 9%

reduction. The mitigation of the L/F losses in the serial system is likely due to task dependencies, wherein worker heterogeneity will invariably lead to system imbalance. Since for a given simulation treatment, we optimize the task scheduling, worker-task rotation through cross training multifunctional workers leads to more balanced system operation. Nonetheless, in practice there may be significant L/F losses from multifunctionality in serial systems.

Effects of Workforce Heterogeneity

Heterogeneity represents the magnitude of individual L/F characteristic differences. Nembhard and Shafer (2007) showed that increasing the level of workforce heterogeneity in parallel systems leads to potential improvements in productivity. A similar result is observed in figure 5.4(right). The current study also extends this finding to the serial system (Figure 5.5 left). Further, all worker selection policies tend to improve productivity with higher levels of workforce heterogeneity (Figure 5.7). When increasing the magnitude of heterogeneity (variance) by a factor of three, the relative improvement in productivity is 20% in the serial case and 17% in the parallel case.

Conclusions

The aim is to provide potential input to managers for the general process of selecting workers for a set of tasks, with the overall goal of maximizing the production. Parallel and series task scenarios were examined to represent a range of configurations, in that many production systems involve a combination of these two types. This study investigates the selection of a group of workers and their assignments to a set of tasks. The selection policies performances are examined with respect to the ratio of specialists to generalists, the level of workforce heterogeneity, and the multi-functionality of the generalists. While the selection process examines four policies relative to a baseline, task assignments are determined optimally based on individual L/F characteristics, in order to gauge the effects of the selection policies.

As a result, selecting workers based on their maximum prior expertise provided the highest throughput in both production scenarios with 34.5% and 28% improvements in parallel and serial systems, respectively. This is confirmatory of earlier findings that higher levels of workforce heterogeneity are associated with higher production levels. While resource flexibility

has been described to be a valuable asset for responding to uncertain demand, full flexibility has also been described as undesirable. The current study agrees with theoretical results from Nembhard and Bentefouet (2012), for parallel systems, in which maximum throughput is achieved through specialization. In serial systems, a small amount of flexibility, by using generalists can improve production, but full flexibility, tends to reverse those gains.

Among the selection policies, opting for a greedy approach that consists of selecting workers based on the highest expected output at any task (policy *MaxiMax O*) performs no better than the baseline policy, while selecting workers based on the highest level of prior expertise at any task leads to best result tested. That is, a greedy policy based on prior expertise is somewhat predictive of general performance, but examining the whole L/F function, tends to negate any gains, due to the nature of the empirical L/F distribution. Selection processes based on previous expertise and the fact that workers will spend time on multiple tasks reinforces the importance of experience. A worker may be less likely to be used specifically for the high output task for which s/he had been selected. The poor performance of the *MaxiMax O* policy represents a hiring or selection process that is not designed in accordance with the ultimate objectives of production.

An extension of the current work is to consider that by working closely with other workers who are doing the same or a related job, a transfer of knowledge is likely. That is, learning in a working place can be the result of learning by doing as well as transfer of knowledge from other individuals. The ultimate goal is to provide managers with decision guidelines that can improve productivity through increased transfer of knowledge among employees learning new skills or receiving cross training.

CHAPTER 6

SCHEDULING WITH LEARNING BY DOING AND LEARNING BY TRANSFER OF KNOWLEDGE WITH HUMAN PERFORMANCE VARIABILITY

6.1 Introduction

In this chapter, in addition to learn by doing, the investigation includes another way of learning: learning from others by transfer of knowledge. The study assesses the importance of acknowledging learning by transfer of knowledge when scheduling workers to tasks. The chapter is organized as follows. Section 6.2 is devoted to the presentation of the learning model that includes both learning by doing and learning by transfer of knowledge. Section 6.3 focuses on the methodology. In Section 6.4, we summarize the results of the simulation design and discuss the outcome in section 6.5.

6.2 A model for learning by doing and learning by transfer of knowledge

Learning by doing has been studied and quantified by several models developed by researchers. Mazur and Hastie (1978) describe a hyperbolic function that behaves well for both mechanical and cognitive learning by doing. The model presented in equation $(6-1)$, was shown to provide the best fit among eleven models for manual skills learning (Nembhard and Uzumeri, 1999). Their multi-criteria comparison considered the average model fit among individuals, the variance of the fit, the number of parameters, and the ability to capture episodes of negative behavior. The hyperbolic function is of the following form:

$$y = k\left(\frac{x+p}{x+p+r}\right)$$

such that $y, k, p, x \geq 0$ and $p + r > 0$ $(6-1)$

where y is a measure of work performance, x is the amount of cumulative work in units of time or trials, and p represents the individual's accumulated prior experience for that task in the same

units as x. The parameter k estimates the asymptotic limit of y and is the steady-state performance level that can be expected when all learning has been completed. Parameter r is the cumulative production required to attain an output level of $k/2$. The three parameters $k, p,$ and r have the advantage to have straightforward interpretation which should aid in the process of transferring knowledge of workplace performance into a sound of managerial decision making process.

Nembhard and Uzumeri (2000a) modified the three parameters hyperbolic learning function in order to include a forgetting effect. Here, we concentrate on learning and propose a modified expression of the productivity rate expressed in equation $(6-1)$. When the operator is working in a team environment, not only does he learn from his own experience but also from the transfer of knowledge taking place from other operators. We consider this learning process by modeling the knowledge of the subject as the sum of his own knowledge obtained by practicing and the knowledge obtained by learning from the other subjects in the same environment. The proposed model has the following form:

$$y_x = k\left(\frac{\theta * T + x + p}{\theta * T + x + p + r}\right) \qquad (6-2)$$

where the additional parameters are: θ the transfer factor and T the total pool of knowledge (cumulative work of all operators). The parameter θ quantifies the percentage of learning that is transferred from other workers. It can represent both negative and positive transfer of knowledge. When the transfer parameter θ is bounded between 0 and 1, there is no unlearning or negative transfer of knowledge to the subject. The model has the advantage to be easily extended to include negative transfer of knowledge by simply allowing negative value for θ. Another advantage of the model presented in equation $(6-2)$ is that each parameter, k, p, r and θ have convenient interpretations, inferences on parameters and relationships between parameters can aid in the sound managerial decision making process, as this would be based on the actual knowledge of workplace performance. For simulation, we substitute to the model presented in equation $(6-2)$ the following expression:

$$y_{ijt} = k_{ij}\left(\frac{\theta_{ij} * T_{ijt} + \sum_{k=1}^{t} x_{ijk} + p_{ij}}{\theta_{ij} * T_{ijt} + \sum_{k=1}^{t} x_{ijk} + p_{ij} + r_{ij}}\right) \qquad (6-3)$$

where i is the worker index, j the task index, and t the index of the production period. The cumulative output x presents in equation $(6-2)$ is replaced by the term $\sum_{k=1}^{t} x_{ijk}$ that cumulates the number of times the worker i has been assigned to task j up to period t. The term T_{ijt} is defined in appendix A and estimates the productivity at time t of other workers performing tasks similar to task j.

6.3 Methodology

The study aims to address important questions managers have to face in the workplace when organizing the production. We conduct several simulations and to acknowledge that decision-making can occur at distinct steps when allocating workers to tasks, we generate a pool of 45 workers and select nine who are then group and later assign to tasks. We therefore organize the scheduling problem into three consecutive stages: 1) a selection stage that involves the question of who; based on a ranking method, workers are selected from the workers population 2) a grouping stage which purpose is to illustrate the transfer phenomenon that can occur when workers are organized in sub-groups and perform the same type of task 3) an assignment stage, the question of what tasks is tackled with sub-group of workers assigned to certain type of tasks. The performance of a worker on a task is estimated using the productivity model described in equation (3). We distinguish two types of parameters: the learning parameters $k, p,$ and r which are generated by a multivariate normal distribution with mean vetor μ and covariance matrix Σ (see table 1), and the transfer parameter θ that is obtained after fitting an univariate normal distribution with mean $m = 0.6641$ and variance $\sigma = 0.4091$.

Table 6. 1. Learning parameter means, variance-covariance matrix

$$\mu = \begin{bmatrix} 1.448 \\ 1.963 \\ 2.0446 \end{bmatrix} \begin{matrix} \log k \\ \log p \\ \log r \end{matrix} \qquad \Sigma = \begin{matrix} \log k \\ \log p \\ \log r \end{matrix} \begin{bmatrix} 0.0045 & 0.0167 & 0.0191 \\ 0.0167 & 0.4621 & 0.2443 \\ 0.0191 & 0.2443 & 0.1905 \end{bmatrix}$$

$$\log k \qquad \log p \qquad \log r$$

The production setting: we consider three types of machine designated as machine 1, machine 2 and machine 3. Each type of machine is replicated three times as a task such that the ultimate number of tasks considered in the experiment is nine. It is assumed that for each instance of a type of machine, a worker has the same learning and transfer parameters. Two

production scenarios are of interest with their respective configuration given in figure 6.1 and figure 6.2.

The selection stage: we examine the question of how to select a workforce for a production system. We consider several rankings method. With each ranking method, the nine top workers are retained for the grouping and assignment stages. A ranking method is based on selecting the top nine workers with either, the highest (*MaxiMax*) or the highest minimum (*MaxiMin*) value with regard to a parameter (learning parameter, transfer parameter or expected output) among all three types of tasks.

The grouping stage: In order to illustrate the importance of the transfer of knowledge that has been discussed to occur when workers perform similar tasks in the same environment, workers selected by a ranking method are organized in teams. The nine selected workers from the selection stage are grouped in exclusive teams with three workers by team. Various strategies are implemented when assembling the teams. A grouping strategy is based on constituting team such that the sum of the teams variance with regard to a parameter (learning parameter, transfer parameter or expected output) on all three types of tasks is either maximal (max variance) or minimal (min variance). We define the variance within a team as the sum of the square variations of the learning parameter, transfer parameter, or expected output between workers of the given team.

The assignment stage: the assignment process addresses the question of which type of task to perform by team of workers. As mentioned above, each team of three workers is assigned to a set of tasks that are replications of a unique type of machine. The workers within a team will only perform tasks which are replications of the type of machine the team has been assigned to. An assignment strategy is based on assigning the teams to the corresponding tasks that has either

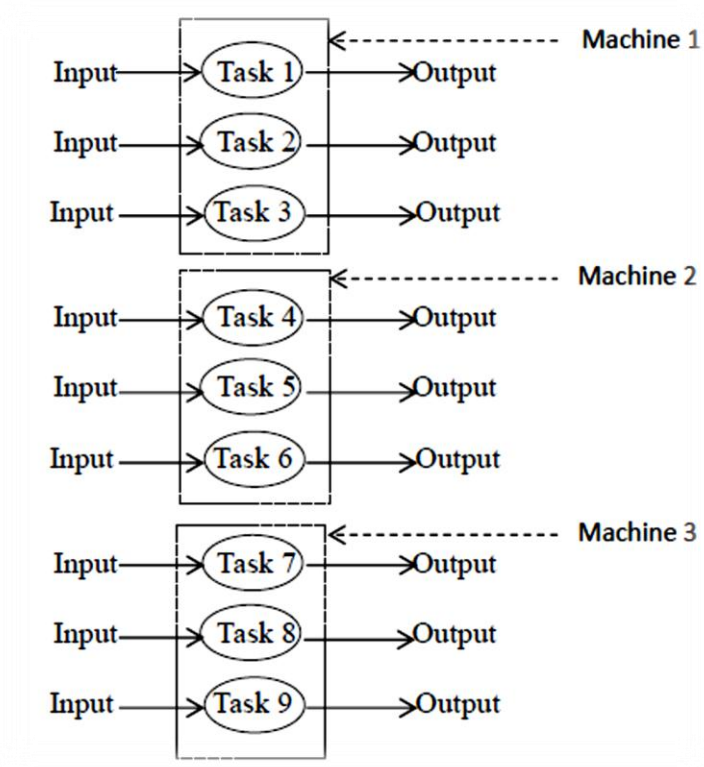


Fig 6. 1. Parallel system configuration (9 tasks)

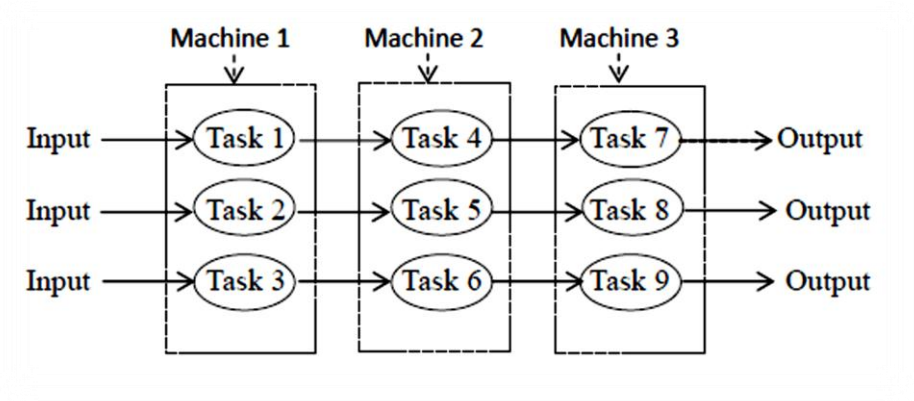


Fig 6. 2. Serial system configuration (9 tasks)

the highest (*MaxiMax* assignment) or the lowest (*MiniMax* assignment) parameter value. The parameter value can be the learning parameter, the transfer parameter, or the expected output. The process starts by assigning the team with the worker that has either, the highest or the lowest parameter or expected output among all three teams. Once the assignment of a team is made, the same assignment procedure goes on with the remaining teams.

The different strategies at the selection, the grouping and the assignment stages are summarized in table 6.2. A solution to the allocation problem is then a combination of a selection method, a grouping objective and an assignment strategy. Given the different objectives at the three different stages, we have a total of one thousand feasible ways to select, group and assign workers. Experiments are then conducted to test and compare the various strategies. An experiment run is: given a population of 45 workers, we apply selection methods based on workers ranking then for each selection method, grouping strategies are implemented and for each grouping strategy, assignment approaches are simulated. We consider 25 runs and for each run we generate a new population of 45 workers.

While the *Maximax* approaches during the selection, grouping and assignment stages can be view as a "greedy" strategy, the *Maximin* approaches tend to have balanced skilled workers. We are interesting in the decision making at different stage of the worker-to-task allocation problem. Among the questions we tackle are; the importance of each stage in the allocation problem, the relevancy to consider worker parameters at various stage of the worker-to-task problem with regard to productivity improvement.

Table 6. 2. Policy by stage by parameter

Stage	**Objective**	**Parameters**
Selection	Maximax	$(O, k, p, 1/r, \theta)$
	Maximin	$(O, k, p, 1/r, \theta)$
Grouping	Max Var	$(O, k, p, 1/r, \theta)$
	Min Var	$(O, k, p, 1/r, \theta)$
Assignment	Maximax	$(O, k, p, 1/r, \theta)$
	Maximin	$(O, k, p, 1/r, \theta)$

O is the expected output with the productivity model described in equation (6-1)

The policies presented in the previous section include a selection stage that starts with a population of 45 workers which is then reduced to nine workers for the grouping and the assignment stages. While being representative of the workplace, the selection, the grouping, and the assignment stages are done separately, which does not guarantee the optimality of the ultimate allocation of worker-to-task. To evaluate the quality of the implemented policies, their

performance are compared to a baseline random allocation of workers to tasks and to an upper bound solution obtained by solving a non-linear scheduling problem where we relax the grouping constraint and the exclusive task type restriction. The mixed integer nonlinear program (MINLP) is presented in appendix A. The baseline random allocation of worker to tasks is obtained by first randomly select 9 workers among the 45 workers, then we arbitrary group them in exclusive teams of 3 workers and finally we randomly assign team of workers to type of task. Each run is replicate 20 times. The following section presents the results of the experiments.

6.4. Result

The ANOVA result of the experiments is shown in table 6.3. All three factors; selection, grouping, and assignment are significant in both production scenarios. The mean square column reveals that the selection stage accounts for a major part of the variability in both parallel and serial system, further followed by the assignment stage and the grouping stage. At the exception of the interaction between the selection and the grouping stage, all two-level interactions are significant in parallel and serial production. The three-level interaction between selection, grouping, and assignment is not significant in either scenario.

Table 6. 3. ANOVA table for simulation treatment

		Parallel output			Serial output		
Source	d.f.	*Mean Sq*	*F-value*	*Pr > F*	*Mean Sq*	*F-value*	*Pr > F*
Selection	9	721716.7	3092.05	0.000*	77058.3	144.6	0.000*
Grouping	9	2294.6	9.83	0.000*	2780.5	5.22	0.000*
Assignment	9	154441.4	661.67	0.000*	29811.2	55.94	0.000*
Selection*Grouping	81	197.4	0.85	0.837	1379.9	2.59	0.000*
Selection*Assignment	81	1170.8	5.02	0.000*	1223.1	2.3	0.000*
Grouping*Assignment	81	8050.9	34.49	0.000*	2743	5.15	0.000*
Selection*Grouping*Assignment	729	241.2	1.03	0.2625	398.3	0.75	1

*Significance at 5% level

Table 6. 4. Top 10 policy parallel system

Selection	Grouping	Assignment	Mean throughput	Mean relative gap to upper bound in % (std in %)
Maximax k	Min variance k	Maximax k	808.89 (3.681)	1.15 (1.03)
Maximax k	Min variance p	Maximax θ	808.89 (3.681)	1.15 (1.03)
Maximax k	Min variance k	Maximax θ	807.77 (3.602)	1.15 (1.06)
Maximax k	Min variance $1/r$	Maximax k	807.77 (3.602)	1.15 (1.06)
Maximax k	Min variance $1/r$	Maximax θ	807.77 (3.602)	1.15 (1.06)
Maximax k	Min variance p	Maximax k	806.76 (4.124)	1.16 (1.06)
Maximax k	Min variance θ	Maximax θ	806.56 (3.818)	1.16 (0.95)
Maximax k	Min variance θ	Maximax k	804.88 (6.148)	1.18 (1.03)
Maximax k	Min variance O	Maximax p	803.02 (14.025)	1.54 (1. 36)
Maximax θ	Min variance O	Maximax p	803.00 (13.715)	1.56 (1.43)

Table 6.4 and table 6.5 display the top ten policies respectively in the parallel and the serial system. A policy is a combination of a selection, grouping, and assignment. The first three columns starting from the left detail the objective at each stage, the fourth column gives the mean performance value as well as the standard deviation over 25 runs. The last column illustrates the efficiency of the considered policy by calculating in percentage the mean and the standard deviation of the relative gap to an upper bound obtained by solving the allocation problem presented in appendix K.

For either parallel and serial scenario, the top ten policies are very closed in term of performance (less than 1% of relative gap between the top and the tenth policy in both systems) and display appreciable efficiency (in average less than 2 % of relative gap to the upper bound for the parallel system, and less than 7 % in the serial system). Selection strategy based on maximax k and maximax θ dominate the top ten policies.

Table 6. 5. Top 10 policy serial system

Selection	Grouping	Assignment	Mean throughput	Mean relative gap to the upper bound in % (Std in %)
Maximax θ	Min variance $1/r$	Maximax O	259.07 (7.670)	6.22 (3.14)
Maximax θ	Min variance k	Maximax θ	258.02 (7.691)	6.28 (3.14)
Maximax θ	Min variance k	Maximax k	258.63 (7.812)	6.38 (3.29)
Maximax k	Min variance O	Maximax O	258.62 (7.409)	6.38 (3.03)
Maximax θ	Max variance p	Maximax k	258.56 (6.421)	6.42 (2.43)
Maximax k	Min variance k	Maximax θ	258.48 (7.275)	6.44 (3.02)
Maximax θ	Max variance p	Maximax θ	258.40 (6.437)	6.47 (2.42)
Maximax k	Min variance p	Maximax θ	258.11 (7.216)	6.57 (2.94)
Maximax O	Min variance k	Maximax k	257.78 (6.994)	6.69 (2. 98)
Maximax k	Min variance O	Maximax $1/r$	257.63 (12.522)	6.75 (4.73)

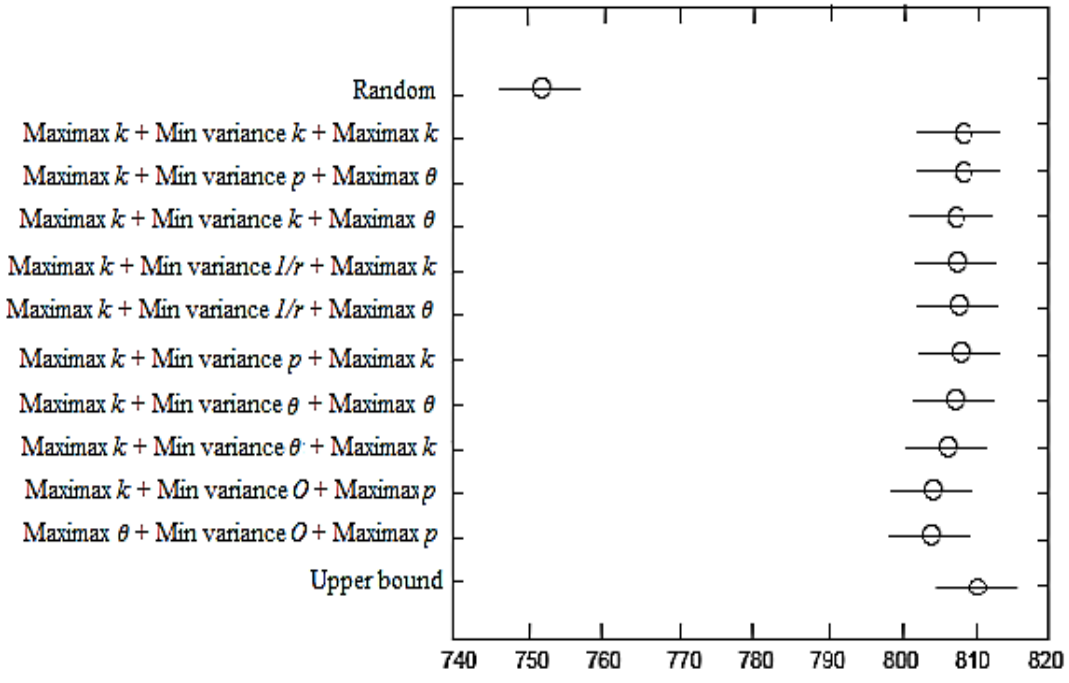


Fig 6. 3. Policies vs Random and Upper bound- Parallel system

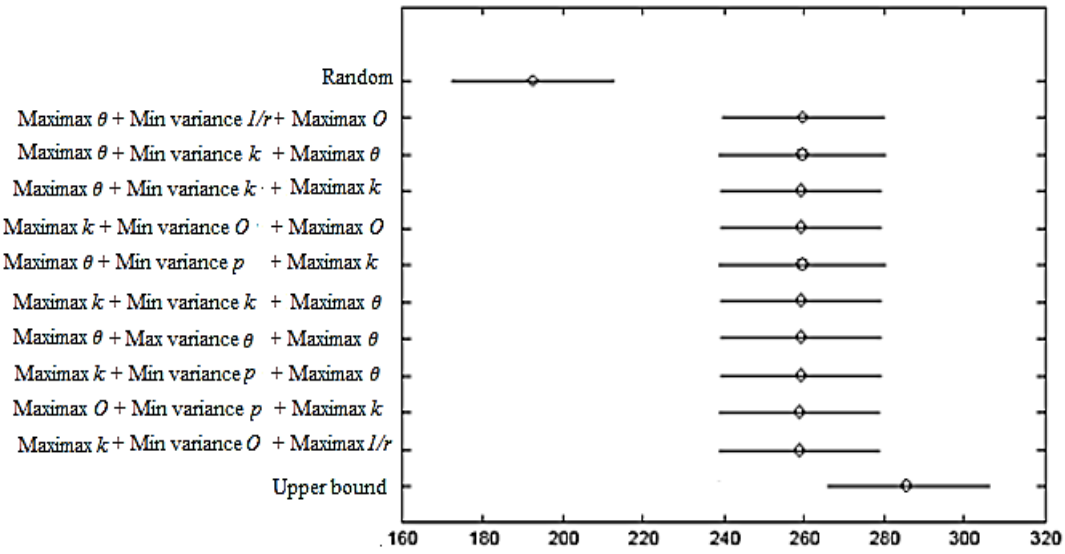


Fig 6. 4. Policies vs Random and Upper bound- Serial system

Figure 6.5 to figure 6.10 display the mean effect for each of the three stages in both production systems. Figure 5 and figure 6 show that in both systems, parallel and serial, at the exception of the parameter r (we consider the parameter $1/r$ rather than r since the parameter r quantifies the number of produced units needed to reach half of the asymptotic steady state productivity), when selecting workers, the maximax policies lead to better system performance compare to the maximin policies. Maximax policies with regard to the output (O), the asymptotic steady state (k), and the previous expertise level (p) are significantly different in the serial system and are not significantly different in the parallel production scenario. Among all policies, selecting workers based on the maximax $1/r$ policy leads to the least production performance.

Figure 6.7 and figure 6.8 illustrate the impact of the implemented policies when grouping workers in teams. In the independent tasks scenario, when assembling workers in teams, the policies that minimize the variance are in average preferable to the ones that maximize the variance. However the difference in term of system throughput between the two approaches, having similar or dissimilar workers within teams, is only significant when considering the parameter expected task output. For the related tasks scenario, the difference between having dissimilar or similar workers with regard to a specific parameter is not significant.

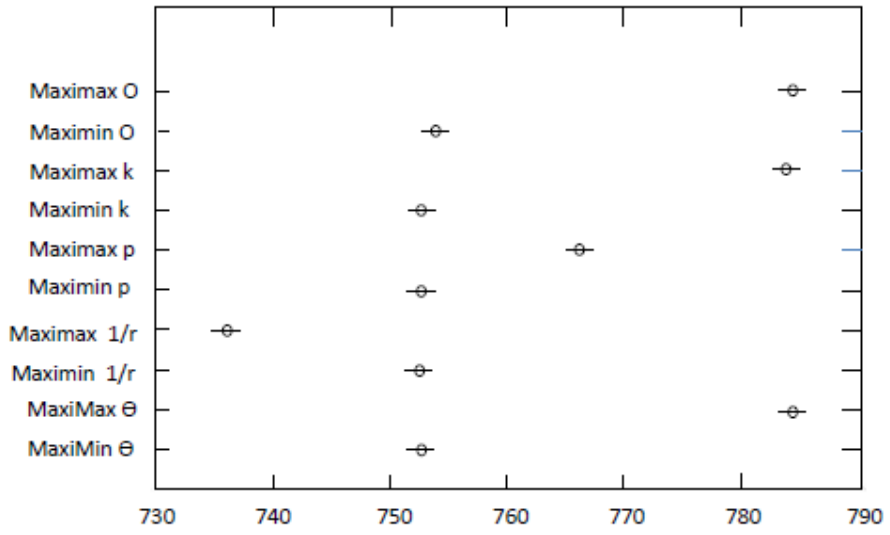


Fig 6. 5. Mean effect of selection policy - Parallel system

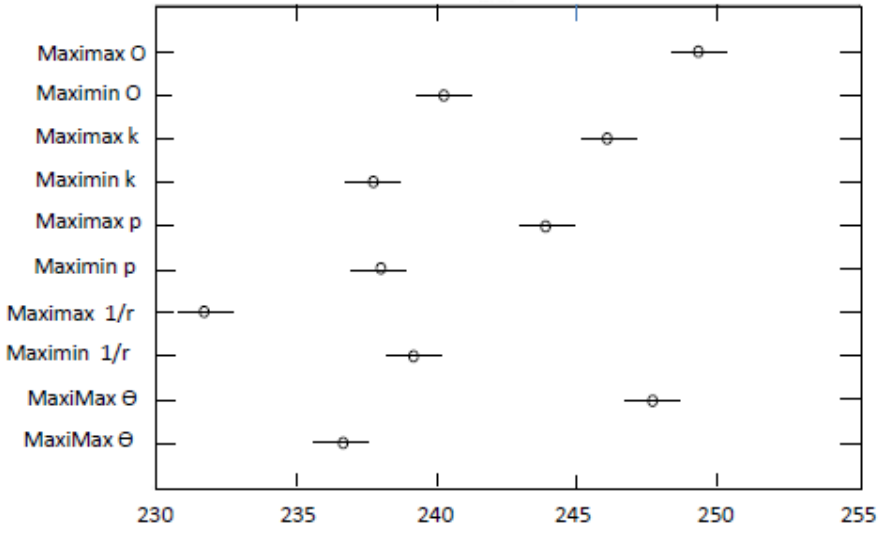


Fig 6. 6. Mean effect of selection policy - Serial system

At the grouping stage, the relative difference in system throughput between the least and the most effective policy among those implemented is less than 0.7 percent in the parallel production scenario and less than 2 percent in the serial production scenario. Figure 7 and figure 8 allow assessing the performance of implemented policies by comparing the system throughput to the upper bound obtained after solving the formulation presented in appendix K.

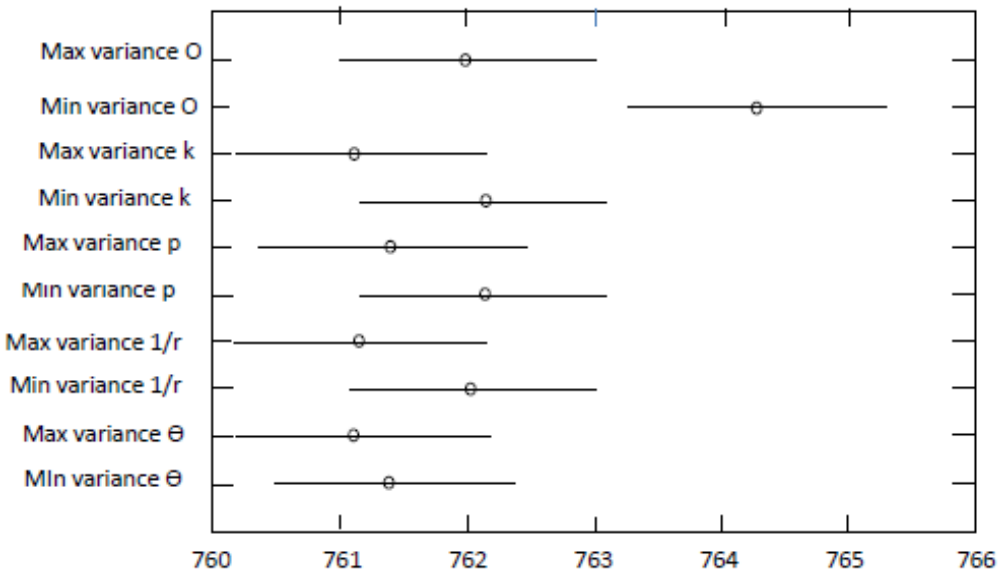


Fig 6. 7. Mean effect of grouping policy - Parallel system

Figure 6.9 and figure 6.10 show the importance of each policy in the process of assigning team of workers to tasks. As in the selection stage in both systems, parallel and serial, at the exception of the parameter*r*, the maximax policies when assigning team of workers to tasks lead to better system performance compare to the maximin policies.

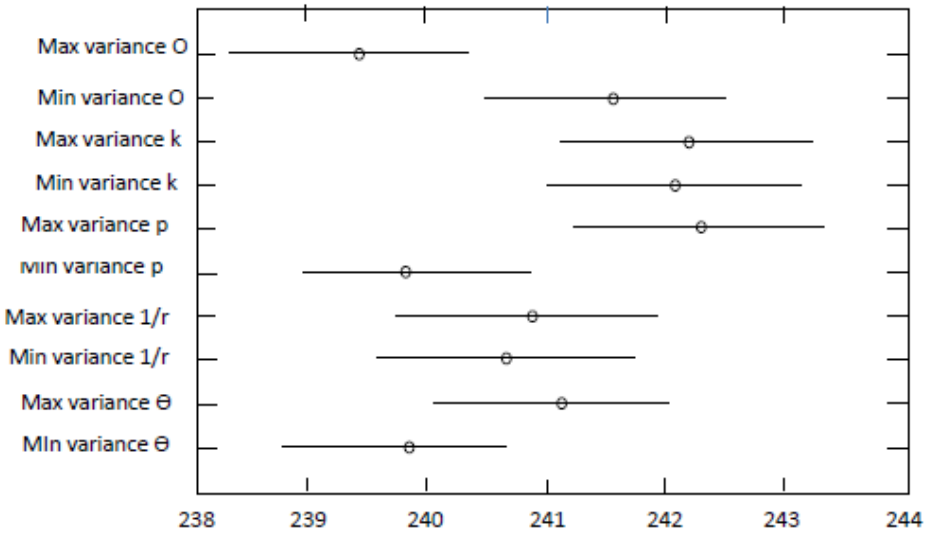


Fig 6. 8. Mean effect of grouping policy - Serial system

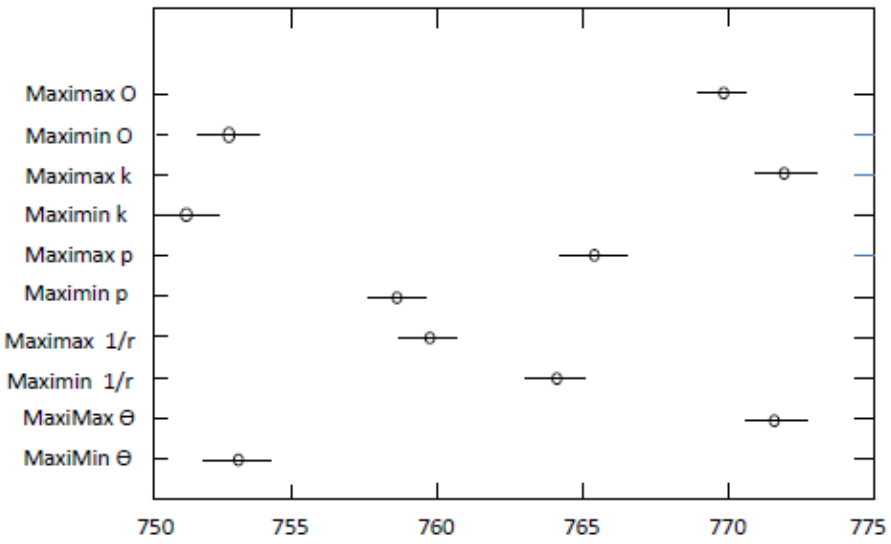


Fig 6. 9. Mean effect of assignment policy - Parallel system

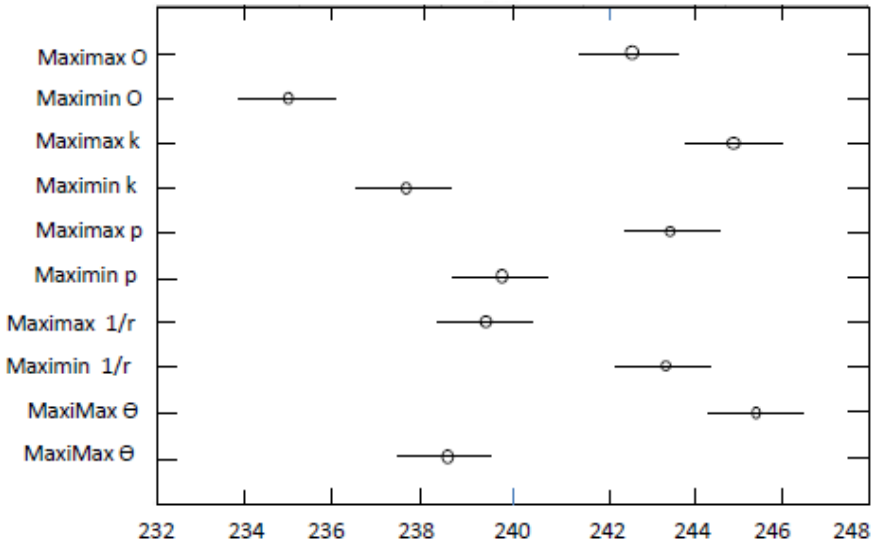


Fig 6. 10. Mean effect of assignment policy - Serial system

Figure 6.11 to figure 6.16 display the two-level interactions between the different stages. Figure 6.11 illustrates the variation of the selection policies with regard to the grouping strategies in the parallel tasks configuration while the case of the serial production scenario is displayed in figure 6.12.

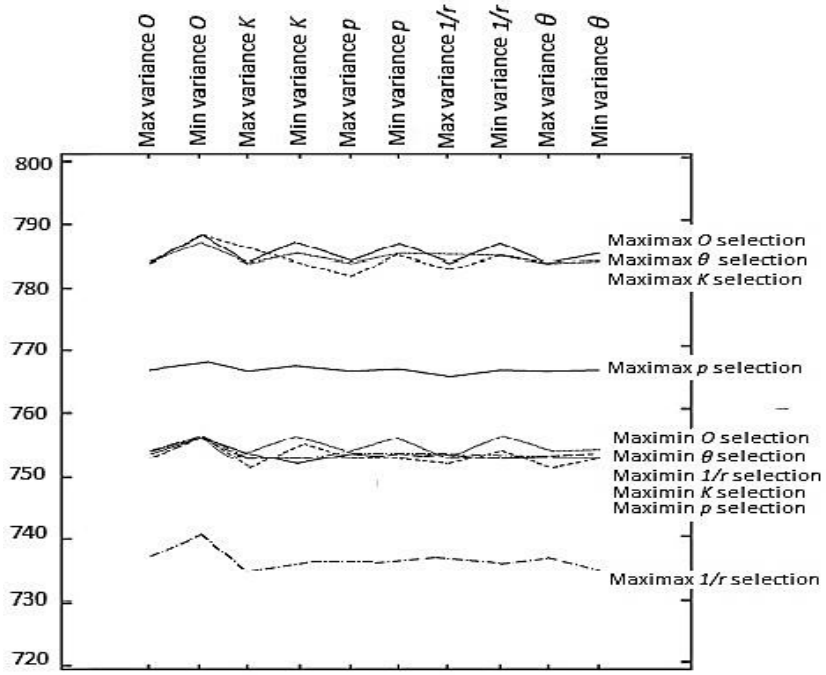


Fig 6. 11. Interaction selection stage and grouping stage - Parallel system

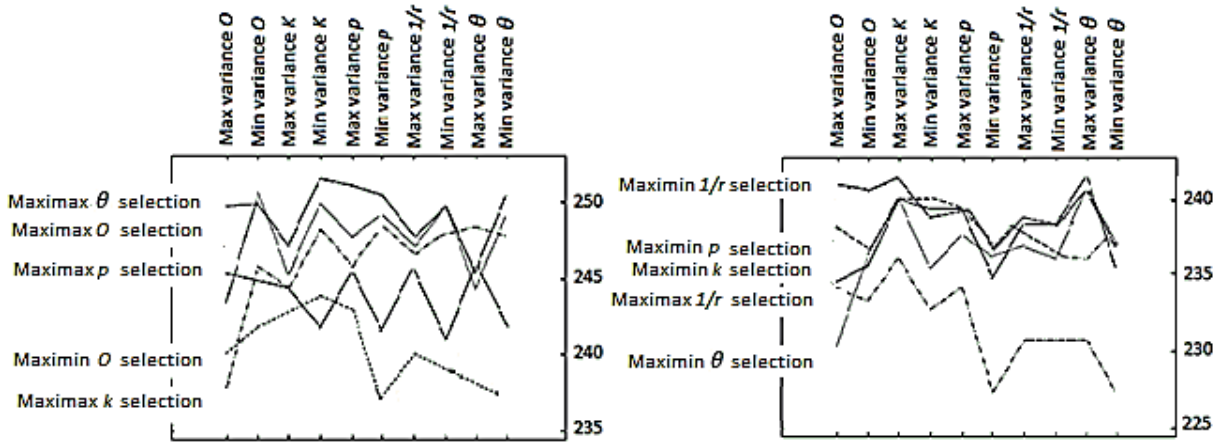


Fig 6. 12. Interaction selection stage and grouping stage - Serial system

In the parallel system, each selection policy shows little variation form one grouping strategy to another. The same observation can be made for the case of the serial configuration even if there is no clear dominance of one selection strategy. We observe that in average the highest level of performance for a policy is not necessary realize by the grouping strategy with regard to the same parameter as in the selection stage. For instance, in the serial configuration in average the maximax θ selection strategy is optimized by adopting a grouping tactic that minimizes team variance with regard to parameter *k*.

Figure 6.13 for parallel tasks and figure 6.14 for serial tasks illustrate the two-level interaction between the two majors source of variability. For either configuration, parallel or serial, each selection tends to have the same saw tooth pattern for the throughput variation with regard to the assignment strategy with the tops associated to the maximax assignment strategies and the lows with the maximin assignment policies.

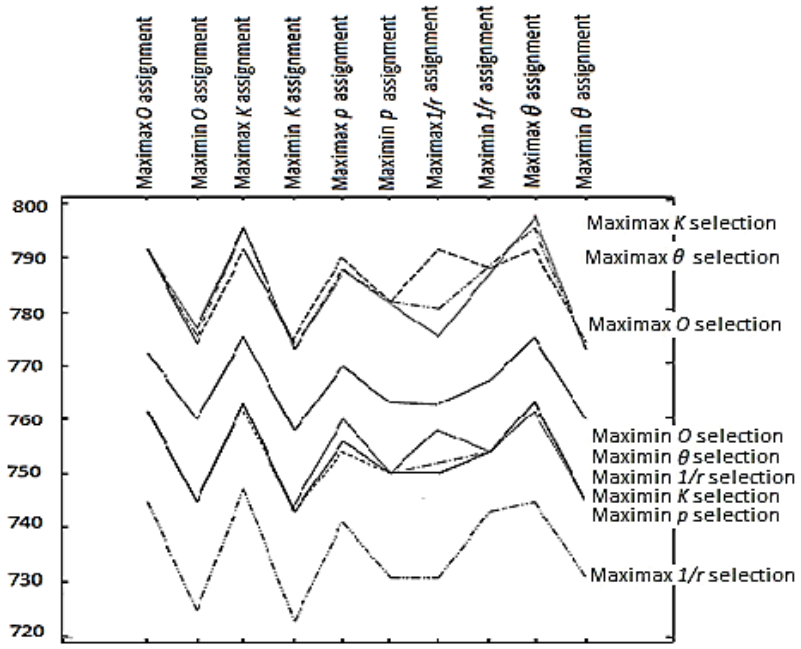


Fig 6. 13. Interaction selection stage and assignment stage - Parallel system

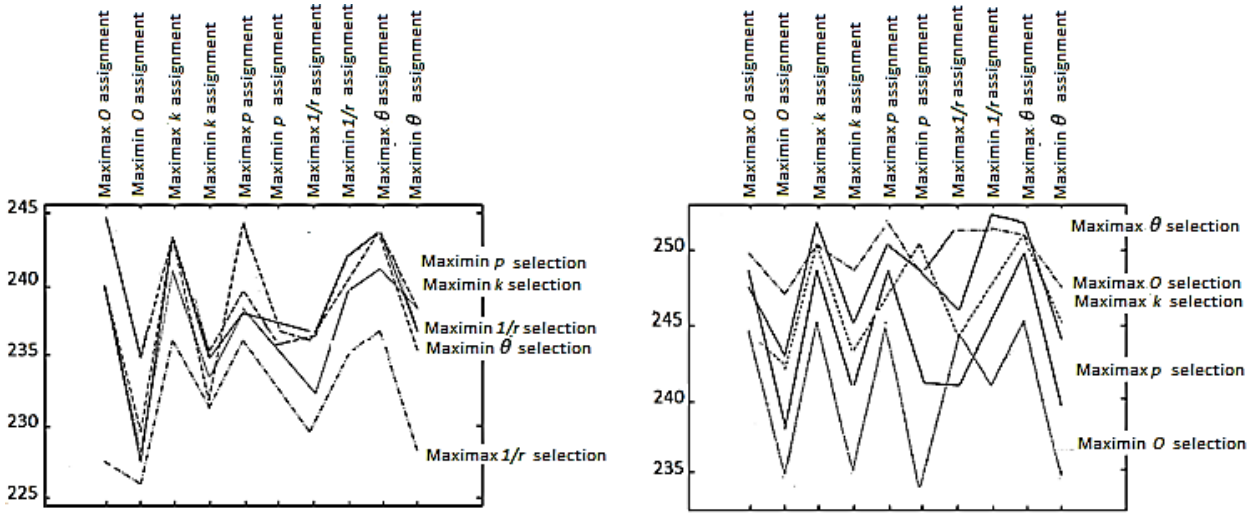


Fig 6. 14. Interaction selection stage and assignment stage - Serial system

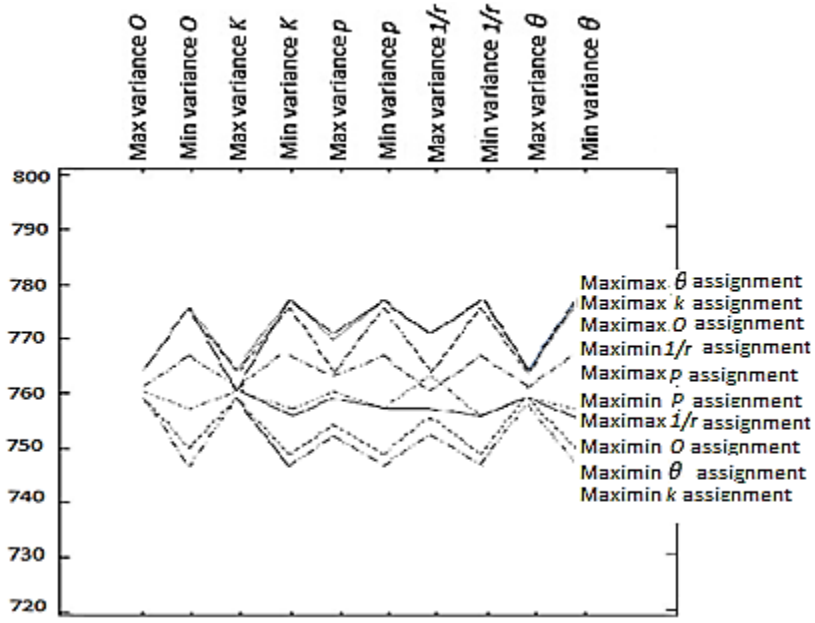


Fig 6. 15. Interaction assignment stage and grouping stage - Parallel system

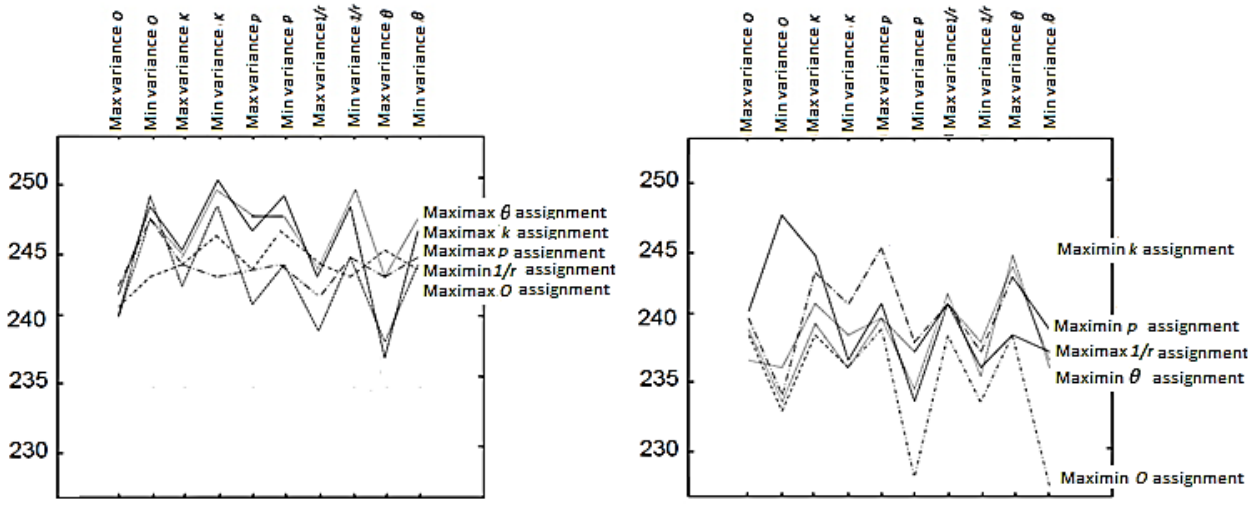


Fig 6. 16. Interaction assignment stage and grouping stage - Serial system

Figure 6.15 and figure 6.16 focus on the interaction between the assignment stage and the grouping stage. We observe the repetition of the saw tooth pattern, but this time the tops are associated with the minimization of the variance and the lows with the maximization of the variance when grouping workers. One can observe that in figure 6.15 in the case of the parallel system, similar to the variation of the selection policies with regard to the grouping strategies, the performance of the assignment policy displays little changes with respect to the grouping tactic. For the serial configuration, while some assignment policy show little changes (maximin

1/r, maximax *p*, and maximin *Θ)* others display significant improvement (maximin *O*, maximax *k*, maximax *1/r*) .

6.5. Discussion

The study investigates the selection and allocation of workers to tasks. In addition to a selection and an assignment stages, a grouping stage is added to assess the importance of learning by transfer. Because we model learning by transfer to occur when workers perform similar tasks, the grouping stage consists of assembling workers in teams which later are assigned to a type of task. To distinguish the two phenomena, learning by doing and learning by transfer, the parameters that account for learning *k, p, r,* are sampled independently of the parameter θ that quantifies the transfer of knowledge. The data used to generate learning by doing and learning by transfer parameters are obtained from the workplace. We consider two systems production: independent and dependent tasks.

From table 6.3 and from figure 6.3 to figure 6.8, the three stages can be ordered by their influence on the system throughput as follow: selection, assignment, and grouping. Table 6.3 illustrates that the major source of variation is due to the selection stage distantly follow by the assignment stage. Compare to the two other stages, the variation due to the grouping stage is negligible which suggests that with regard to policies at the selection and the assignment phases, having similar or dissimilar workers in teams has a low impact on the system throughput. At the selection stage, the top nine workers with regard to the considered parameter are selected. As a consequence of that strategy in the selection and because of the final size of the workforce (nine workers are hired), we end up hiring the workers whose considered parameter value (k, p, r, and θ) are closed to each other that is the selection process by design prevents important differences between the workers who will be grouped.

Among the implemented policies at the selection stage, the maximax policies with regard to parameter O, k and θ dominate the top 30 best policies as displayed in appendix B. The parameter O is obtained by calculating with equation $(6-1)$ the output when performing a task during the same duration as for the experiment hence it gives a fairly good estimate of the worker production at the end of the simulation. The performance of the policies with regard to k is explained by the parameter's role in the hyperbolic model. Nembhard and Shafer (2007) study

in detail the hyperbolic model. The experiment duration is equivalent to a production length of one year that allows the selected workers to reach their asymptotic steady state productivity. When the steady state productivity is reached at a task, the productivity model becomes a linear function of the parameter *k* then the performance of policy with regard to *O* becomes similar to the one with regard to *k*. The parameter θ measures the amount of knowledge that can be transferred to a worker from other workers performing a similar task. The transferred knowledge is materialized by an increment in the productivity rate, increment proportional to the production of workers performing similar tasks. The performance of maximax θ selection policy suggests that selecting workers with regard to the ability to learn from co-workers is as important as hiring workers based on their learning by doing parameters. Therefore at the recruitment stage, an attention should be given to test the ability of a worker to learn from others.

As underlined in the previous paragraph, the various considered selection policies undermine the importance of the grouping stage. Hence, grouping workers by similarities or dissimilarities with regard to a parameter does not have an important impact on the system throughput with the considered selection and assignment policies. The assignment stage accounts for the second factor in explaining variation of the system throughput. The same explanation that accounts for the dominance of policies based on parameters O, k and θ at the selection stage holds for their dominance at the assignment stage. The policy performance at the assignment stage for the parallel system displays results similar to the ones at the selection stage with the top policies being maximax policies with regard to *O*, *k* and θ. At the assignment stage, with the maximax policies the best worker among each team best worker with regard to the considered parameter decides which task the team he belongs will be assigned to, once complete for one team, the same methodology is applied to the remained not assigned teams and so on. Since the parameter *O* is obtained by estimating the output of a worker at each task during a period that equals the simulation duration and because specialization is optimal in parallel system (Nembhard and Bentefouet, 2012a), it is expecting that a maximin assignment policy with regard to *O* performs well. The simulation length allows workers to reach the steady state productivity hence the maximax assignment policy with regard to *k* is expected to have a performance in the same order as the maximax assignment policy with regard to *O*. For the same reason as in the selection stage, the performance of the maximax assignment policy with regard to the transfer

parameter θ should also performs well. In a serial system as illustrate by Bentefouet and Nembhard (2012b), the throughput of the production line is the output from the least productive task of the line. The performance of a line depends on all workers assigned to any task of that line hence the gap observed in the parallel system between the best policies and the others is likely to be reduced in the serial system in order to account for these dependencies.

The top ten heuristic policies performance are compared to a random policy and an upper bound obtained by solving the formulation presented in appendix A. Heuristic based on parameter O, k, and θ for the selection stage provides the best performance in the independent task scenario as the relative gap between the upper bound and the respective heuristics is around 2%. Especially in the parallel system, the combination of the parameter O being obtained by estimating worker performance at task during the simulation duration, workers reaching their asymptotic steady state productivity and workers learning form co-workers explain the performance of the policies. The absence of the maximin selection strategies, is explained by the fact that maximin policies favor workers that are efficient at all tasks of the workplace. However with the experiment setting, workers are assigned to one type of task hence workers selected on their ability to perform well at each of the nine tasks are not as productive at the unique task they are finally assigned to compare to the workers selected on their performance at a single type of tasks which they are ultimately assigned.

Conclusion

The study seeks to provide inputs to managers for the general process of selecting, grouping and assigning workers to a set of tasks. Rather than rely on heavy nonlinear optimization problem that can be overwhelming to implement and to solve, we try to build straightforward policies. Parallel and series task scenarios were examined to represent a range of configurations and since many production systems involve a combination of these two types. The study investigates the management of workers in the context of performance variability. We extend the original model from Mazur and Hastie (1978). The productivity function considered includes learning effect. The proposed model accounts for learning by doing and learning by transfer of knowledge from co-workers performing similar duties. For the purpose of studying transfer of knowledge, we consider three stages in the allocation of workers to tasks; we select workers, then group them in exclusive teams, and finally assign each team to a unique type of tasks. The implemented

policies rely on learning parameters of the hyperbolic three factors model and a transfer of knowledge parameter. We examine the policy performance at each stage of the worker-to-task problem. We compare their performance to a baseline random allocation of workers to tasks and to an upper bound established by solving a math program optimization problem in which we ignore the restriction to have workers in teams and teams to be assigned to exclusive tasks.

The investigation reveals that; when hiring top workers with regard to their ability to perform the tasks (output, learning by doing parameters) or to learn from co-workers (transfer parameter), assembling them in teams of similar or dissimilar workers has a lower impact on the production performance compare to the selection and the assignment strategies. Our experiment results suggest that selecting, grouping and assigning workers based either on their estimated output at each task, the asymptotic steady state or their ability to learn from other workers provide the best performance in a parallel system. In the serial system the dependencies between tasks make difficult to back one heuristic against another. All heuristics provides better performance in the serial system compare to a random policy. It is also showed that top policies constitute relevant heuristics since the gap with the upper bound, obtained by relaxing the teams grouping and the restriction on tasks to perform is in the range of 2% in a parallel system and a range of 7 % in a serial system.

Scheduling workers based on the agility to learn on practice have been discussed to improve system performance. Here we develop a model that also accounts for the transfer of knowledge. The investigation shows that scheduling with regard to the ability to learn from others is as relevant as scheduling with regard to learning by doing. An extension of the current work is to include workforce learning effect in traditional supply chain optimal ordering policies. It is often assumes that facilities have steady productivity rate however one can speculate that the learning effect that occurs at the individual level ultimately impact the overall performance of the organization. It is uncertain to which level, collective individual worker performance improvement significantly impact supply chain delivery.

CHAPTER 7

CONCLUSION AND FUTURE RESEARCH

The presented work bridges the gap between the classic theory of scheduling and human performance literature. The research addresses the worker-to-task allocation problem under human performance variability. Insights on optimal scheduling policies are provided and heuristics with performance close to optimal are suggested. In the context of this document, human variability has been modeled to occur with two phenomena: learning and forgetting. The two effects differentiate human from machine and influence the worker performance on a task. The presented research has established the following results:

- Assigning workers to single task during the production time is an optimal policy in parallel system when the objective is to maximize throughput.
- In a parallel system configuration with demand requirements at each task, the allocation problem under general productivity function can be formulated as a linear math program and it is optimal to limit the number of stint per worker per task to a maximum of one.
- When maximizing throughput in serial system, depending on workers performance at tasks, work sharing or specialization is an optimal policy when scheduling workers to tasks. Solving the serial system scheduling problem is equivalent to maximize the least productive task and optimal switching time can be characterized.
- Under workforce heterogeneity and multi-functionality, with the workforce composed of specialist workers and generalist workers, and when learning and forgetting is included, selecting workers on prior expertise outperformed greedy selection approach. Full flexibility and specialization policies are not optimal policies in serial production scenario.
- When considering learning by transfer of knowledge and learning by doing in parallel and serial production scenarios, selecting, grouping, and assigning workers based on learning and transfer parameters are heuristics that display performance close to optimal policy. When hiring top workers with regard to their ability to perform tasks (output,

learning by doing parameters) or to learn from co-workers (transfer parameter), assembling them in teams of similar or dissimilar workers has a lower impact on the production performance compare to the selection and the assignment strategies.

The dissertation concludes by pointing to potential directions for future research. The thesis focuses on the organization at the factory level. Within the factory, reducing the task completion times and building schedules that integrate characteristics of the resources would lead to a better system performance. However, when integrated in a supply chain, schedules have to be elaborate in accordance with the flow of goods and resources. This requires a well coordinate supply chain from suppliers to retailers. In the past, organizations have focused their efforts on taking efficient decisions at each stage of the supply chain.

Recently, there has been a trend of moving from decoupled decision making processes toward more coordinated and integrated design and control of all stages from the supply to the delivery of goods and services to the customers. Managing a supply chain often implies synchronizing material flows between suppliers and customers. It has been discussed that coordination between stages is one of the main issues of supply chain management. For instance in customer driven supply chain it is said that procurement and production policies should be coordinated. The importance of purchased materials costs – In high technology firms, it accounts for up to 80% of the total product cost – has put more emphasis on supplier's selection and inventory control. The coordination between the replenishment and the inventory policy is an area of research where more applicable integrated models can be developed. While it is clear that having the right set of skilled workers influences the productivity and ultimately the delivery time to the customer, most of the attention in supply chain management has been devoted to the flow of products and less on the process of selecting, assigning, and ultimately scheduling heterogeneous workers to a set of locations. One of the research questions is to connect the workforce skills matrix to the supply chain objectives.

Appendix A. Proof of Theorem 3.1

Consider the formulation ***P***, with a monotone *learning* function $f_{ij}\,(t)$, that is non-decreasing and continuous in time as described by Badiru (1992), and Anzanello and Fogliatto (2007). To represent the output for a period in a continuous form, it is assumed that the beginning of the period *t* coincides with the end of period *t-1*. This argument is supported by the fact that forgetting is excluded (For clarity, the case with forgetting is examined separately). Hence, the level of learning accumulates from one period to the next. Nembhard and Shafer (2007) use the same integral representation to consider the output through a given time period, wherein:

$$O_{i,j,t} = \int_{t-1}^{t-1+x_{i,j,t}} f_{ij}\,(u - \tau_{ijt})\,du \qquad (\mathrm{A}-1)$$

Hence, $O_{i,j,t}$ is the output of worker *i* on task *j* *during* period *t*, and depends on the cumulative experience of worker *i* gained on task *j* prior to period *t*, that is, $\sum_{k=1}^{t-1} x_{i,j,k}$.

In Equation (A-1), τ_{ijt} is the cumulative number of periods the worker has not been assigned to the task prior to time *t*, whereby the production during period *t* is observed at the end of the period.

$$\tau_{ijt} = (t-1) - \sum_{k=1}^{t-1} x_{i,j,k} \qquad (\mathrm{A}-2)$$

The output produced during period *t* in Equation (A-1) is found via an integral with a lower limit equal to $t-1$, and upper limit of $t-1\ + x_{i,j,t}$. These limits of integration imply that the output at period *t* is equal to the production between the end of period $t-1$ and the end of period $t-1\ + x_{i,j,t}$. That is, if $x_{i,j,t} = 0$, worker *i* is not assigned to task *j* during period *t*, hence the output from worker *i* on task *j* during period *t* is null. If $x_{i,j,t} = 1$, worker *i* is assigned to task *j* during period *t*, hence the output from worker *i* on task *j* during period *t* depends on the worker skills and the knowledge acquired prior to period *t*. The productivity function f_{ij} in Equation (A-1) is evaluated at $u - \tau_{ijt}$, where *u* belongs to the interval defined by the limits of integration. At

time *t*, the cumulative experience gained by worker *i* is $t - \tau_{ijt}$ rather than *t*. In order to simplify the expression of the integrand in Equation (A-1), we define

$v = u - \tau_{ijt}$ and recall the definition of τ_{ijt} given by Equation (A-2), resulting in Equation (A-3).

$$O_{i,j,t} = \int_{\sum_{k=1}^{t-1} x_{i,j,k}}^{\sum_{k=1}^{t} x_{i,j,k}} f_{ij}\,(v)dv \qquad (\mathrm{A}-3)$$

Letting $n_{ijt} = \sum_{k=1}^{t} x_{i,j,k}$ be the number of periods worker *i* is assigned to the task *j* up to a given time *t*, it is straightforward to note that, n_{ijt} = 0, and Equations (A-4) and (A-5) hold by the definitions of n_{ijt}, and *T*.

$$\sum_{i=1}^{I} n_{ijT} = T \qquad \forall j \qquad (\mathrm{A}-4)$$

$$\sum_{j=1}^{J} n_{ijT} = T \qquad \forall i \qquad (\mathrm{A}-5)$$

Equation (A-4) holds, since having an idle station and idle worker during any period will be sub-optimal, as an idle worker and station can be paired to increase total production. Equation (A-5) holds in a similar fashion, since there is an equal number of tasks and workers. By substituting Equation (A-4), Equation (A-3) can be rewritten as Equation (A-6).

$$O_{i,j,t} = \int_{n_{ijt-1}}^{n_{ijt}} f_{ij}\,(v)dv \qquad (\mathrm{A}-6)$$

Let $F_{i,j,T}$ describe the *cumulative* production of worker *i* on task *j* during time horizon *T*. That is, $F_{i,j,T}$ sums the output from each period *t, t = 1,..., T*, and from Equation (A-6), one can quantify the cumulative production as in Equation (A-7).

$$F_{i,j,T} = \sum_{t=1}^{T} \int_{n_{ijt-1}}^{n_{ijt}} f_{ij}\,(v)dv = \int_{0}^{n_{ijT}} f_{ij}(v)dv \qquad (A-7)$$

First consider the case of two workers, two tasks and *T* periods of production,

The total production on a single task *j* over time horizon *T* from both workers is given in Equation (A-8).

$$\sum_{\mathrm{i}=1}^{2} F_{i,j,T} = \int_0^{n_{1jT}} f_{1j}(v)dv + \int_0^{n_{2jT}} f_{2j}(v)dv \qquad (A-8)$$

Recalling Equation (A-4), Equation (A-8) can be rewritten in terms of a single n_{ijT} variable with respect to a worker. That is, knowing the number of assignments of a given worker provides information about the other worker.

$$\sum_{\mathrm{i}=1}^{2} F_{i,j,T} = \int_0^{n_{1jT}} f_{1j}(v)dv + \int_0^{T-n_{1jT}} f_{2j}(v)dv \qquad (A-9)$$

Thus, Equation (A-10) gives the overall *system production* (two workers, two tasks and *T* periods) as the sum of production over the two tasks, applying Equation (A-9) twice.

$$\sum_{\mathrm{j}=1}^{2}\sum_{\mathrm{i}=1}^{2} F_{i,j,T} = \int_0^{n_{11T}} f_{11}(v)dv + \int_0^{T-n_{11T}} f_{21}(v)dv + \int_0^{T-n_{11T}} f_{12}(v)dv + \int_0^{n_{11T}} f_{22}(v)dv \qquad (A-10)$$

The productivity function f_{ij} is a twice differentiable monotonically non-decreasing function over time. Thus, by defining F_{ij} as an anti derivative of f_{ij}, for $t \geq 0$ we have Equation (A-11).

$$F_{ij}(t) = \int_0^t f_{ij}(v)dv \qquad (A-11)$$

Common properties of this anti-derivative may be stated as follows.

$$\frac{dF_{ij}(t)}{dx} = f_{ij}(t) \;\; and \;\; \frac{dF_{ij}(T-t)}{dx} = -f_{ij}(T-t) \quad t \in [0,T].$$

Recalling that f_{ij} is a twice differentiable monotonically non-decreasing function over time that is $F_{ij}(t)$ and $F_{ij}(T-t)$ are convex functions (see e.g. Rockafellar, 1970). Thus, Equation (A-10) may be rewritten as in Equation (A-12)

$$\sum_{\mathrm{j}=1}^{2}\sum_{\mathrm{i}=1}^{2}F_{i,j,T}=F_{11}(n_{11T})+F_{21}(T-n_{11T})+F_{12}(T-n_{11T})+F_{22}(n_{11T}) \quad (A-12)$$

From Equation (A-12), it follows that the system production $\sum_{\mathrm{j}=1}^{2}\sum_{\mathrm{i}=1}^{2}F_{i,j,T}$ is a function of a single variable n_{11T} (the number of periods worker 1 is assigned to task 1), since *T* is assumed to be a given. It is also a convex function as it is the sum of convex functions. Thus, maximizing the overall system production and solving formulation ***P*** for two workers and two tasks is equivalent to solving Equation (A-13)

$$\max_{t\,\epsilon\,\{0,T\}}\sum_{\mathrm{j}=1}^{2}\sum_{\mathrm{i}=1}^{2}F_{i,j,T} \quad (A-13)$$

Recalling $F_{i,j,T}=\sum_{t=1}^{T}O_{i,j,t}$, Equation (A-13) stipulates that the maximum total production is achieved at either *t = 0* or *t = T*. These two options correspond to schedules where the workers are specialized during the time horizon. Thus, in the case of two workers and two tasks with a general model of learning specialization is optimal.

Extending the above argument to the general case of *I* workers and *J* tasks, where we let *m=I=J*, during *T* periods of production, and recalling Equation (A-4), the production on task *j* is a function of *m-1* variables$\left(n_{ij}\right)_{1\le i\le m-1}$ the output of task *j* from all workers during *T* may be written as in Equation (A-14), where $\sum_{\mathrm{i}=1}^{\mathrm{m}}F_{i,j,T}$ is a convex function of each independent variable $n_{ij}, 1\le i\le m-1$.

$$\sum_{\mathrm{i}=1}^{\mathrm{m}}F_{i,j,T}=$$
$$\int_0^{n_{1j}}f_{1j}(v)dv+\int_0^{n_{2j}}f_{2j}(v)dv+\cdots+\int_0^{n_{m-1j}}f_{m-1j}(v)dv+\int_0^{T-\sum_{i=1}^{m-1}n_{ij}}f_{mj}(v)dv \quad (A-14)$$

It follows that the system production $\sum_{j=1}^{m}\sum_{\mathrm{i}=1}^{\mathrm{m}}F_{i,j,T}$ with *m* tasks and *m* workers is a function of *m-1* vectors $\boldsymbol{n_i}=(n_{i1},\dots,n_{im-1})^t$. For notational simplification, we will denote F_T as the system production of the *m* tasks and *m* workers, where F_T is the function of vectors $(\boldsymbol{n_i})_{1\le i\le m-1}$ given in Equation (A-15).

$$F_T=\sum_{j=1}^{m}\sum_{\mathrm{i}=1}^{\mathrm{m}}F_{i,j,T}\left(n_{1j},\dots,n_{m-1j}\right) \quad (\mathrm{A}-15)$$

Combining Equations (A-14) and (A-15) yields an expression for the system production during over the time horizon in Equation (A-16).

$$F_T = \sum_{j=1}^{m} \left(\int_0^{n_{1j}} f_{1j}(v)dv \; + \cdots + \int_0^{n_{m-1j}} f_{m-1j}(v)dv \right.$$
$$\left. + \int_0^{T-\sum_{i=1}^{m-1} n_{ij}} f_{mj}(v)dv \right) \qquad (A-16)$$

Taking the partial derivative of continuous differentiable function F_T with respect to n_{ij} yields Equation (A-17), which holds for $1 \leq i, j \leq m-1$, and is the sum of non-decreasing functions of the variable n_{ij}. Thus, Equation (A-17) is a convex function of each argument n_{ij}.

$$\frac{\partial F_T}{\partial n_{ij}} = f_{ij}(n_{ij}) - f_{ij}\left(T - \sum_{j=1}^{m-1} n_{ij}\right) - f_{mj}\left(T - \sum_{i=1}^{m-1} n_{ij}\right) \qquad (A-17)$$

Returning to the original optimization problem, ***P***, we can expand F_T to see that the convexity property again leads to the simpler problem as Equation (A-18).

$$\max_{0 \leq n_{ij} \leq T, 1 \leq i,j \leq m-1} F_T \left[\begin{pmatrix} n_{11} \\ \vdots \\ n_{1j} \\ \vdots \\ n_{1\,m-1} \end{pmatrix}, \ldots, \begin{pmatrix} n_{i1} \\ \vdots \\ n_{ij} \\ \vdots \\ n_{i\,m-1} \end{pmatrix}, \ldots, \begin{pmatrix} n_{m-11} \\ \vdots \\ n_{m-1j} \\ \vdots \\ n_{m-1\,m-1} \end{pmatrix} \right]$$
$$= \max_{n_{ij} \in \{0,T\},\ 1 \leq i,j \leq m-1} F_T \qquad (A-18)$$

Hence, the optimal schedule for a system of independent tasks and independent jobs with *m* workers who learn, *m* tasks and *T* periods of production is a schedule with workers specialized each to a single task.

Appendix B. Proof of Theorem 3.2

Equation (A-6) in the proof of Theorem 3.1, expresses the production of worker *i* on task *j* during period *t.* It follows that the total production of worker *i* on task *j* for the full time horizon *T* is the sum of the outputs of each period as in Equation $(B\ -\ 1)$.

$$\sum_{\mathrm{t=1}}^{\mathrm{T}} O_{i,j,t} = \sum_{t=1}^{T} \int_{n_{ijt-1}}^{n_{ijt}} f_{ij}\,(v)dv$$

$$= \int_{0}^{n_{ijT}} f_{ij}\,(v)dv \qquad (B-1)$$

Given the production time horizon *T*, we let F_{ij} be the total production of worker *i* on task *j,* whereby F_{ij} is a function of a unique variable n_{ijT}, the cumulative time worker *i* was assigned to task *j*, or one assignment of worker *i* to task *j* with length n_{ijT}. That is, a worker will not return to a given task, but may work for a length of time less than T at that task. Equation $(B-1)$ suggests that instead of having non-contiguous stints of worker-task assignments, one can always frame it as one stint with length equal to the total cumulative duration of the time it spent on that specific task. Thus, there exists an optimal solution to the parallel-task scheduling problem with the maximum number of stints of a worker on a task equal to one when there are demand requirements to meet during the time horizon.

Appendix C. Proof of Lemma 4.1

The MaxMin formulation presented in section 4.2.1 of chapter 4 optimizes allocation of worker to tasks in order to maximize the output from the bottleneck station. It is a relaxation of the general flow-line problem, since buffer constraints are ignored. Hence, the optimal value of the objective Q in section 4.2.1 is an upper bound of the flow line's maximum throughput.

Consider two workers indexed by i_p, $p \in \{1,2\}$ and two tasks indexed by j_q, $q \in \{1,2\}$, let denote $m^* = \left(m^*_{i_p j_q}\right)_{p,q}$ an optimal time allocation of workers to tasks when solving the MaxMin problem (formulation *Q*). $m^*_{i_p j_q}$, the time spent by worker *i* at task *j* is defined by:

$$m^*_{i_p j_q} = \sum_{t=0}^{T} m_{i_p j_q\, t} * t \qquad (C-1)$$

$m_{i_p j_q\, t}$ is a binary decision variable of the MaxMin problem that is 1 if worker i_p is assigned to task j_q for a duration t. Given m^* from equation $(C-1)$, the goal is to build a feasible schedule to the flow line problem that has no more than one change point (switch of workers). With two workers and two tasks, there are only two assignments possible: The first, A_1 worker i_1 at task j_1 and worker i_2 at task j_2 , and the second, A_2 worker i_1 at task j_2 and worker i_2 at task j_1. When allowing a maximum of one change point, the total output within the flow line is:

$$\begin{cases} P = min\left(F_{i_1 j_1}, F_{i_2 j_1}\right) + min\left(B + F_{i_2 j_1}, F_{i_1 j_2}\right), \quad B = max\left(F_{i_1 j_1} - F_{i_2 j_2}, 0\right) \\ \text{If started with } A_1 \\ \text{or} \\ P = min\left(F_{i_2 j_1}, F_{i_1 j_2}\right) + min\left(B + F_{i_1 j_1}, F_{i_2 j_2}\right), B = max\left(F_{i_2 j_1} - F_{i_1 j_2}, 0\right) \\ \text{If started with } A_2 \end{cases} \qquad (C-2)$$

$F_{i_p j_q}$, B and P are respectively the production from worker i_p on task j_q, the buffer level and the total production from the assembly line. Let denote $\left(F^*_{i_p j_q}\right)_{p,q}$ the output obtained from m^*,

the MaxMin optimal time allocation and one can build a flow line feasible solution whose structure is defined by equation $(C-2)$.

a. $F^*_{i_1 j_1} \geq F^*_{i_2 j_2}$

Consider the schedule that starts with assignment A_1 for a lapse $t_1 = max(m^*_{i_1 j_1}, m^*_{i_2 j_2})$ and then switch to assignment A_2 for $t_2 = max(m^*_{i_2 j_1}, m^*_{i_1 j_2})$. The total production is:

$$P = min(F^*_{i_1 j_1}, F^*_{i_2 j_2}) + min(max(F^*_{i_1 j_1} - F^*_{i_2 j_2}, 0) + F^*_{i_2 j_1}, F^*_{i_1 j_2})$$

$$= F^*_{i_2 j_2} + min(F^*_{i_1 j_1} - F^*_{i_2 j_2} + F^*_{i_2 j_1}, F^*_{i_1 j_2})$$

$$= min(F^*_{i_1 j_1} + F^*_{i_2 j_1}, F^*_{i_1 j_2} + F^*_{i_2 j_2}) = min(F^*_{j_1}, F^*_{j_2})$$

That is, the flow line throughput equals the output of the least productive station, which is the MaxMin objective.

b. $F^*_{i_1 j_1} \leq F^*_{i_2 j_2}$ and $F^*_{i_2 j_1} \leq F^*_{i_1 j_2}$

Consider the schedule staring with assignment A_2 during $t_1 = max(m^*_{i_2 j_1}, m^*_{i_1 j_2})$ and switching to assignment A_1 during $t_2 = max(m^*_{i_1 j_1}, m^*_{i_2 j_2})$. The total production is:

$$P = min(F^*_{i_2 j_1}, F^*_{i_1 j_2}) + min(max(F^*_{i_2 j_1} - F^*_{i_1 j_2}, 0) + F^*_{i_1 j_1}, F^*_{i_2 j_2})$$

$$= F^*_{21} + min(F^*_{i_1 j_1}, F^*_{i_2 j_2})$$

$$= F^*_{i_1 j_1} + F^*_{i_2 j_1}$$

$$= F^*_{j_1}$$

Under the assumptions $F^*_{i_1 j_1} \leq F^*_{i_2 j_2}$ and $F^*_{i_2 j_1} \leq F^*_{i_1 j_2}$ the least productive task is task j_1, that is we again have that the flow line throughput equals the MaxMin objective.

c. $F^*_{i_1 j_1} \leq F^*_{i_2 j_2}$ *and* $F^*_{i_2 j_1} \geq F^*_{i_1 j_2}$

Consider the schedule that begins with assignment A_2 during $t_1 = max(m^*_{i_2 j_1}, m^*_{i_1 j_2})$ and then apply assignment A_1 during $t_2 = max(m^*_{i_1 j_1}, m^*_{i_2 j_2})$. The total production is:

$$P = min\left(F^*_{i_2 j_1},\ F^*_{i_1 j_2} \right) + min\left(max\left(F^*_{i_2 j_1} - F^*_{i_1 j_2}, 0 \right) + F^*_{i_1 j_1},\ F^*_{i_2 j_2} \right)$$

$$= F^*_{i_1 j_2} + min\left(F^*_{i_2 j_1} - F^*_{i_1 j_2} + F^*_{i_1 j_1},\ F^*_{i_2 j_2} \right)$$

$$= min\left(F^*_{i_1 j_1} + F^*_{i_2 j_1},\ F^*_{i_1 j_2} + F^*_{i_2 j_2} \right) = min\left(F^*_{j_1},\ F^*_{j_2} \right)$$

Similar to the outcome of (a) we state that the flow line throughput equals the MaxMin objective. Since subsections (a), (b), and (c) consider all possible scenarios that is in the case of two workers and two tasks, solving the MaxMin problem is equivalent to solving the flow line problem.

Appendix D. Proof of Theorem 4.1

Denote $f_{i_p\, j_q}$, the steady state productivity of worker i_p on task j_q, $q \in \{1,2\}$ $p \in \{1,2\}$ and introduce the following notation: $\alpha = \frac{f_{i_2\, j_1}}{f_{i_1\, j_1}}$, $\delta = \frac{f_{i_2\, j_2}}{f_{i_1\, j_2}}$, $\gamma = \frac{f_{i_1\, j_1}}{f_{i_1\, j_2}}$ and $\beta = \frac{t}{T}$ is the proportion of the time horizon before the switch occurs. Therefore, the MaxMin formulation introduced in section 3.2.1 is reduced to:

$$Maximize\ Q$$

s.t

$$Q \leq f_{i_1\, j_2} \times T\{\gamma\beta(1-\alpha) + \alpha\gamma\}$$

$$Q \leq f_{i_1\, j_2} \times T\{\beta(\delta - 1) + 1\}$$

$$0 \leq \beta \leq 1$$

(i) $(1-\delta)(1-\alpha) > 0$

The maximum of Q is reached at the intersection of the two lines with slopes of opposite sign defined by the equations $y = \gamma\beta(1-\alpha) + \alpha\gamma$ and $y = \beta(\delta - 1) + 1$.

If $(1-\delta) > 0$, and $(1-\alpha) > 0$, or $(1-\delta) < 0$ and $(1-\alpha) < 0$ provided that $\alpha < \frac{1}{\gamma} < \frac{1}{\delta}$, or $\alpha > \frac{1}{\gamma} > \frac{1}{\delta}$, respectively, the optimal switching point is defined by $t^* = \beta \times T$, where β is given by:

$$\beta = \frac{1 - \alpha\gamma}{\gamma(1-\alpha) + 1 - \delta} \qquad (D-1)$$

The condition on the ordering of α, $\frac{1}{\gamma}$, and $\frac{1}{\delta}$ ensures that the switch point lies in the interval, $[0, T]$ otherwise specialization is optimal.

(ii) $(\delta - 1)(1-\alpha) < 0$

Equations $y = \gamma\beta(1-\alpha) + \alpha\gamma$ and $y = \beta(\delta - 1) + 1$ have slopes with the same sign and the maximum is reached when $\beta = 0$ or $\beta = 1$. In this case it is optimal to have specialized workers.

(iii) $(\delta - 1)(1 - \alpha) = 0$

If $\alpha = 1$ and $\delta = 1$, Equations $y = \gamma\beta(1 - \alpha) + \alpha\gamma$ and $y = \beta(\delta - 1) + 1$ are parallel straight lines, then there is no intersection point and specialization is optimal.

If $\alpha \neq 1$ and $\delta = 1$ (or $\alpha = 1$ and $\delta \neq 1$) provided $\alpha > \frac{1}{\gamma} > 1$ or $\alpha < \frac{1}{\gamma} < 1$ (or $1 > \gamma > \delta$ or $1 < \gamma < \delta$) , the optimal switching point is defined by $t^* = \beta \times T$, where β is given by equation $(B - 1)$.

When relaxing the assumption of theorem 3.1 and allowing each worker to have the *same* non-null constant productivity when performing the two tasks $f_{i_1 j_1} = f_{i_1 j_2} = f_{i_1}$ and $f_{i_2 j_1} = f_{i_2 j_2} = f_{i_2}$, we again obtain cases (i) and (ii). With the previous notation we have $\alpha = \delta, \gamma = 1$ and $(\delta - 1)(1 - \alpha) = (1 - \alpha)^2$ and it is optimal to switch. Since equation $(B - 2)$ becomes $\beta = \frac{1}{2}$ it is always optimal to switch at time $/2$.

Appendix E. Proof of Theorem 4.2

Let the two workers be indexed by i_1 and i_2 and the two tasks, j_1 and j_2. Recalling the MaxMin problem formulation in section 4.2.1, we define the functions H, G on $[0,T]$ by:

$$\begin{cases} H(t) = F_{i_1\, j_1}(t) + F_{i_2\ j_1}(T-t) \\ G(t) = F_{i_2\ j_2}(t) + F_{i_1\, j_2}(T-t) \end{cases} \qquad (\mathrm{E}-1)$$

Assuming that $f_{p\,q}$ is a general productivity rate model and a continuous non-decreasing function of time we have:

$$\begin{cases} \frac{dH}{dt} = f_{i_1\, j_1}(t) - f_{i_2\ j_1}(T-t) \\ \frac{dG}{dt} = f_{i_2\ j_2}(t) - f_{i_1\, j_2}(T-t) \end{cases} \qquad (\mathrm{E}-2)$$

The mild assumption that $f_{p\,q}$ is a differentiable non-decreasing function of t leads to G and H being two differentiable functions with non-decreasing first derivative, and correspondingly convex functions. Denote $X(H,G) = \{t\,\epsilon\,[0,T] |\, H(t) = G(t)\}$ the set of points such that H equals G and consider the function $l(t)$ and its anti derivative $L(t)$ by:

$$\begin{cases} L(t) = \int_0^t l(u)du \\ l(t) = H(t) - G(t) \end{cases} \quad t\epsilon\,[0,T]$$

L is a twice differentiable function of t. Consider three distinct crossing points $t_{q-1}, t_q,$ and t_{q+1} between H and G and suppose we have:

$$\begin{cases} H(t) \geq G(t)\ \ for\ \ t_{q-1} \leq t\ \leq t_q \\ H(t) \geq G(t)\ \ for\ \ t_q \leq t\ \leq t_{q+1} \end{cases}$$

That is, $H(t) \geq G(t)\ \ for\ \ t_{q-1} \leq t\ \leq t_{q+1}$, and the production, $min(H(t), G(t))$, is a convex function on the interval $[t_{q-1}, t_q]$, hence the maximum is reached either at t_{q-1} or t_q. Therefore, in order to maximize production only the crossing point at which the sign of $H - G$ changes is relevant. These specific crossing points are inflection points of the function L. That is the number

of inflection points of L provide an upper limit on the optimal number of switching points. An inflection point t of L, when it is defined, satisfies:

$$\frac{d^2L(t)}{dt^2} = 0 \Leftrightarrow \left[f_{i_1\, j_1}(t) + f_{i_1\, j_2}(T-t)\right] - \left[f_{i_2\ j_2}(t) + f_{i_2\ j_1}(T-t)\right] = 0 \qquad (\mathrm{E}-3)$$

Appendix F. Proof of Corollary 3.2

When using the hyperbolic model from Mazur and Hastie (1978), the productivity function satisfies:

$$f_{i_p\, j_q}(u) = k_{i_p\, j_q}\frac{t + p_{i_p\, j_q}}{t + p_{i_p\, j_q} + r_{i_p\, j_q}} \quad p \in \{1,2\} \quad q \in \{1,2\}$$

Recalling equation $(\mathrm{C}-3)$, the necessary condition on a switching point is equivalent to solve:

$$\alpha t^4 + \beta t^3 + \gamma t^2 + \delta t + \theta = 0 \qquad (F-1)$$

Where

$$\alpha = k_{i_2\ j_2} + k_{i_2\ j_1} - k_{i_1\, j_1} - k_{i_1\, j_2},\ \beta = k_{i_1\, j_1}A_1 + k_{i_1\, j_2}A_2 - k_{i_2\ j_2}A_3 - k_{i_2\ j_1}A_4$$

$$\delta = k_{i_1\, j_1}C_1 + k_{i_1\, j_2}C_2 - k_{i_2\ j_2}C_3 - k_{i_2\ j_1}C_4,\ \theta = k_{i_1\, j_1}D_1 + k_{i_1\, j_2}D_2 - k_{i_2\ j_2}D_3 - k_{i_2\ j_1}D_4$$

$$A_i = -c_i - a_i\ ,\ B_i = -d_i + a_i c_i - b_i \qquad i = 1, \dots, 4$$

$$C_i = a_i d_i + b_i c_i\ ,\ \ D_i = b_i d_i \qquad i = 1, \dots, 4$$

$$a_1 = T + p_{i_1\, j_2} + r_{i_1\, j_2} - p_{i_1\, j_1},\ b_1 = p_{i_1\, j_1}\left(T + p_{i_1\, j_2} + r_{i_1\, j_2}\right),\ c_1 = T + p_{i_2\ j_1} + r_{i_2\ j_1} -$$

$$p_{i_2\ j_2} - r_{i_2\ j_2},\ d_1 = \left(p_{i_2\ j_2} + r_{i_2\ j_2}\right)\left(T + p_{i_2\ j_1} + r_{i_2\ j_1}\right),\ a_2 = -p_{i_1\, j_1} - r_{i_1\, j_1} + T + p_{i_1\, j_2},$$

$$b_2 = \left(T + p_{i_1\, j_2}\right)\left(p_{i_1\, j_1} + r_{i_1\, j_1}\right),\ c_2 = T + p_{i_2\ j_1} + r_{i_2\ j_1} - p_{i_2\ j_2} - r_{i_2\ j_2},$$

$$d_2 = \left(p_{i_2\ j_2} + r_{i_2\ j_2}\right)\left(T + p_{i_2\ j_1} + r_{i_2\ j_1}\right),\ a_3 = T + p_{i_2\ j_1} + r_{i_2\ j_1} - p_{i_2\ j_2},$$

$$b_3 = p_{i_2\ j_2}\left(T + p_{i_2\ j_1} + r_{i_2\ j_1}\right),\ c_3 = T + p_{i_1\, j_2} + r_{i_2\ j_1} - p_{i_1\, j_1} - r_{i_1\, j_1},$$

$$d_3 = \left(p_{i_1\, j_1} + r_{i_1\, j_1}\right)\left(T + p_{i_1\, j_2} + r_{i_1\, j_2}\right),\ a_4 = -p_{i_2\ j_2} - r_{i_2\ j_2} + T + p_{i_2\ j_1},$$

$$b_4 = \left(T + p_{i_2\ j_1}\right)\left(p_{i_2\ j_2} + r_{i_2\ j_2}\right),\ c_4 = T + p_{i_1 j_2} + r_{i_1 j_2} - p_{i_1 j_1} - r_{i_1 j_1},$$

$$d_4 = \left(p_{i_1 j_1} + r_{i_1 j_1}\right)\left(T + p_{i_1 j_2} + r_{i_1 j_2}\right)$$

The polynomial expression in $(D-2)$ is of degree four, hence we have a maximum of four roots that lead to potential switching points. One can rewrite the polynomial expression in equation $(D-2)$ by denoting it $V(t)$.

$$\begin{cases} V(t) = a_0 t^4 + 4a_1 t^3 + 6a_2 t^2 + 4a_3 t + a_4 \\ if\ \ \alpha > 0, a_0 = \alpha, a_1 = \frac{1}{4}\beta, a_2 = \frac{1}{6}\gamma, a_3 = \frac{1}{4}\delta, a_4 = \theta \\ if\ \ \alpha < 0, \quad a_0 = -\alpha, a_1 = -\frac{1}{4}\beta, a_2 = -\frac{1}{6}\gamma, a_3 = -\frac{1}{4}\delta, a_4 = -\theta \end{cases}$$

Wang and Qi (2005) provide necessary and sufficient conditions for the non-existence of real roots of the quartic $V(t)$ when $a_0 \neq 0$. The non-existence of real roots for the quartic implies that the MaxMin objective function reaches its optimum with specialized workers. When $\alpha = 0$, Kavinoky and Thoo (2007) provide conditions on the existence of real roots for a cubic polynomial function of one variable.

Appendix G. Proof of Corollary 4.3

Consider that each worker has the same productivity at both tasks and knowledge is not transferable from one task to another. When the performance model includes learning, the functions H and G from equation $(G-1)$ in appendix C are reduced to:

$$\begin{cases} H\,(t) = F_{i_1}(t) + F_{i_2}(T-t) \\ \quad and \\ G(t) = F_{i_2}(t) + F_{i_1}(T-t) \end{cases} \qquad (G-1)$$

When a work sharing policy is optimal, the switching point is a crossing point between H and G. From equation $(E-1)$, we have

$$H\,(t) = G\,(T-t) \qquad (G-2)$$

That is, it is only needed to consider the crossing points between H and G that lie in $[0, \frac{T}{2}]$. The assumption on the productivity model remains the same as in appendix C, functions F and G are convex function over $[0, T]$, then they are also convex function over $[0, T/2]$. From equation $(E-2)$, $\frac{T}{2}$ is a crossing point. Because of the convexity of function H and G, any other crossing point between 0 and $\frac{T}{2}$ is dominated (in terms of the objective value Q) either by the objective value with a switch at $\frac{T}{2}$ or a schedule with no switches. By symmetry, we the result is extended to the interval $[0, T]$, and only two policies could be optimal, either specialization or work sharing with a switch at $T/2$. This leads to the following decision rule.

If $H\left(\frac{T}{2}\right) > min\{H\,(T), G\,(T)\}$, it is optimal to switch at $T/2$.

If $H\left(\frac{T}{2}\right) < min\{H\,(T), G\,(T)\}$ then it is optimal to specialize.

If $H\left(\frac{T}{2}\right) = min\{H\,(T), G\,(T)\}$ then it is equivalent to switch at $\frac{T}{2}$, or to specialize.

When relaxing the model by assuming that each worker is indifferent to the two tasks in terms of skill, and knowledge is transferable from one task to another, the set of constraints from the MaxMin formulation is replaced by:

$$Q \leq F_{i_1}(t) + \left[F_{i_2}(T) - F_{i_2}(t)\right]$$

$$Q \leq F_{i_2}(t) + \left[F_{i_1}(T) - F_{i_1}(t)\right]$$

Functions H and G are redefined by:

$$\begin{cases} H(t) = F_{i_1}(t) + \left(F_{i_2}(T) - F_{i_2}(t)\right) \\ \text{and} \\ G(t) = F_{i_2}(t) + \left(F_{i_1}(T) - F_{i_1}(t)\right) \end{cases} \quad (G-4)$$

Equation$(E-4)$ leads to:

$$\begin{cases} \dfrac{dH(t)}{dt} = -\dfrac{dG(t)}{dt} = f_{i_1}(t) - f_{i_2}(t) \\ H(T) = G(0), H(0) = G(T) \end{cases} \quad (G-5)$$

If $f_{i_1} \geq f_{i_2}$ or $f_{i_1} \leq f_{i_2}$, it is optimal to switch, and the switching point is the unique crossing point of H and G.

If f_{i_1} and f_{i_2} cross each other, the set $\{t \in [0,T] \text{ such that } H(t) = G(t)\}$ constitutes the list of optimal switching points.

Appendix H. Proof of Theorem 4.4

With stochastic performance, the objective becomes the maximization of the expected throughput. The problem can be formulated as:

$$Maximize\ Q$$

s.t

$$Q \ \leq \ \mathbb{E}\left[\frac{1}{\tau_{i_1 j_1}}\right] \times t + \mathbb{E}\left[\frac{1}{\tau_{i_2 j_1}}\right] \times (T - t)$$

$$Q \ \leq \ \mathbb{E}\left[\frac{1}{\tau_{i_2 j_2}}\right] \times t + \mathbb{E}\left[\frac{1}{\tau_{i_1 j_2}}\right] \times (T - t)$$

$$0 \leq t \ \leq T$$

$\tau_{i_p j_q}$, the processing time of worker i_p on task j_q, follows an exponential distribution with parameter $\lambda_{i_p j_q}$. The corresponding processing rate, $\frac{1}{\tau_{i_p j_q}}$, has a cumulative density function $F_{\tau_{i_p j_q}}(x) = \begin{cases} e^{-\frac{\lambda_{i_p j_q}}{x}} & , \ x > 0 \\ 0 & , \ x \leq 0 \end{cases}$ and an expected value $\lambda_{i_p j_q} A,$ with $A = \int_0^{+\infty} \frac{1}{z} e^{-\frac{1}{z}} dz$. The set of constraints of the above formulation becomes:

$$\begin{cases} Q/A \ \leq \lambda_{i_1 j_1}\left(1 - \dfrac{\lambda_{i_2 j_1}}{\lambda_{i_1 j_1}}\right) t + \dfrac{\lambda_{i_2 j_1}}{\lambda_{i_1 j_1}} T \\ Q/A \ \leq \ \lambda_{i_1 j_2}\left(\dfrac{\lambda_{i_2 j_2}}{\lambda_{i_1 j_2}} - 1\right) t + \lambda_{i_1 j_2} T \\ 0 \leq t \ \leq T \end{cases} \Leftrightarrow \begin{cases} P \ \leq \lambda_{i_1 j_1}\left(1 - \dfrac{\lambda_{i_2 j_1}}{\lambda_{i_1 j_1}}\right) t + \dfrac{\lambda_{i_2 j_1}}{\lambda_{i_1 j_1}} T \\ P \leq \ \lambda_{i_1 j_2}\left(\dfrac{\lambda_{i_2 j_2}}{\lambda_{i_1 j_2}} - 1\right) t + \lambda_{i_1 j_2} T \\ 0 \leq t \ \leq T \end{cases}$$

The formulation is analog to the one with steady productivity presented in appendix B, with $f_{i_p j_q}$ replaced by $\lambda_{i_p j_q}$.

Appendix I. Proof of Theorem 4.4

Without loss of generality one can assume that for a flow line with J tasks with all jobs having the same sequence, the defined sequence is $1 \to 2 \to \cdots \to J$. By induction, for $J = 1$, the system is a parallel system with one station. Nembhard and Bentefouet (2012) show that the optimal schedule in *parallel* systems with J tasks and I workers with general model of learning is a schedule with specialized workers. The result implies no change points; that is less or equal to zero.

For $= 2$:

If $I \leq 2$, one has a flow line with two workers and two tasks which, by Lemma 3.1 is equivalent to the $MaxMin$ problem. That is, the maximum number of change points is 1.

If $I = 3$, consider the $MaxMin$ formulation with 3 workers and 2 tasks, the set of worker's indexes is $I = \{i_1, i_2, i_3\}$, the set of task's indexes is $J = \{j_1, j_2\}$. Let denote n_{pq}, the time workload of worker i_p at task j_q, $p \in \{1,2,3\}, q \in \{1,2\}$. The feasible region of the MaxMin formulation can be written as:

$$An \leq b \text{ , with } A = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}, n = \begin{bmatrix} n_{i_1 j_1} \\ n_{i_1 j_2} \\ n_{i_2 j_1} \\ n_{i_2 j_2} \\ n_{i_3 j_1} \\ n_{i_3 j_2} \end{bmatrix}, b = T \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

The formulation of the $MaxMin$ problem is a linear program with optimum at an extreme point. Since $rank(A) = 4$, an extreme point is characterized by the intersection of four linearly independent hyperplanes enveloping the feasible region. Four of the five inequalities must be binding suggesting that one worker will be unassigned. That is, solving the *MaxMin* formulation with the objective of maximizing the production in a flow line with two tasks and three workers is equivalent to finding the best combination of at most two workers and solving the *MaxMin* formulation with just two workers and two tasks. Hence, when $I = 3$, the maximum number of change points is 1.

For a general number of workers $I > 2$, the same reasoning can be used to show that $I - 2$ workers are never assigned, and the flow line problem with two tasks and I workers is equivalent to finding the best combination of two workers at most one change point.

Let introduce a concatenation operator, $\cup$, for ordered sets. Consider a J-tasks flow line $[1 \to 2 \to \cdots \to J-1 \to J]$, which can be viewed as a concatenation of subsystems; $[1 \to 2 \to \cdots \to j-1 \to j] \cup [j \to j+1 \to \cdots \to J-1 \to J]$, or the concatenation of $J-1$ subsystems each consisting of two tasks: $[1 \to 2] \cup [2 \to 3] \cup \ldots \cup [j \to j+1] \cup \ldots \cup [J-2 \to J-1] \cup [J-1 \to J]$. These configurations have in common the same minimal required number of change points to reach optimality. Denote $\mathcal{C}(S)$ the minimal number of change points required solving the flow line scheduling problem to optimality with the objective of maximizing the output of the last station:

$$\mathcal{C}([1 \to 2 \to \cdots \to J-1 \to J]) = \mathcal{C}\left(\bigcup_{\mathrm{j}=1}^{\mathrm{j}=\mathrm{J}-1} [j \to j+1]\right)$$

Here, let propose and legitimate the inequality in Equation G-1. Consider 2 systems *A* and *B*, with at least two tasks in each and note that:

$$\mathcal{C}(A \cup B) \leq \mathcal{C}(A) + \mathcal{C}(B) \qquad (\mathrm{I}-1)$$

When combining two systems, it is straightforward to lower the overall number of change points of the combined system by having at least two change points of each system happening at the same time. Iterating inequality $(\mathrm{G}-1)$ leads to:

$$\mathcal{C}\left(\bigcup_{\mathrm{j}=1}^{\mathrm{j}=\mathrm{J}-1} [j \to j+1]\right) \leq \sum_{i=1}^{J-1} \mathcal{C}([j \to j+1]) \qquad (\mathrm{I}-2)$$

Since the required number of change points depends only on the number of tasks involved we have:

$$\mathcal{C}([j \to j+1]) = \mathcal{C}([1 \to 2]) = 1 \quad \text{for all } j \qquad (\mathrm{I}-3)$$

Applying the result of equation $(\mathrm{G}-3)$ to the inequality $(\mathrm{G}-2)$ leads to:

$$\mathcal{C}\left(\bigcup_{\mathrm{j}=1}^{\mathrm{j}=\mathrm{J}-1}[j \rightarrow j+1]\right) \leq \sum_{i=1}^{J-1} \mathcal{C}([1 \rightarrow 2]) = \mathcal{C}([1 \rightarrow 2]) \times \left(\sum_{i=1}^{J-1} 1\right) = 1 \times (J-1)$$

Therefore, $\mathcal{C}([1 \rightarrow 2 \rightarrow \cdots \rightarrow J-1 \rightarrow J]) \leq J-1$. That is, the required number of change points for a J-station system is no greater than $J-1$. This result is independent of the number of workers, *I*.

To motivate the limiting case, consider a system of J stations and an equivalent number of workers. Among the J workers, $J-1$ are relatively inefficient at all J tasks (productivity almost equals to zero) and one worker is highly efficient in all J tasks. The efficient worker must rotate through all stations. This reduces to the *MaxMin* problem with J stations and one worker.

Appendix J. Proof of Theorem 3.5

By induction, Lemma 4.1 shows that in the case of two tasks, solving the flow line is equivalent to solving the *MaxMin* problem. Recalling the three-task scenario from section 4.2.5, the flow line can be viewed as a production system in configuration *A*. By applying the two task result, solving the flow line in configuration *A* (or, *B)* is equivalent to solving the *MaxMin* formulation with the objective of maximizing the least productive *sub-system*:

$$Maximize\ Q$$

s.t

$$Q \leq F_{A_1}$$

$$Q \leq F_{A_2}$$

F_{A_q} , $q = 1, 2$ represents the throughput from the group of tasks A_q at the end of the time horizon *T*.

Task goup A_1 is composed of two tasks tasks, j_1 and j_2, and $F_{A_1} = \min(F_{j_1}, F_{j_2})$. A_2 is reduced to task j_3 and $F_{A_2} = F_{j_3}$. When replacing F_{A_1} and F_{A_2} by the expression as function of tasks j_1, j_2, and j_3 the set of constraints in the above formulation becomes:

$$\begin{cases} Q \leq \min(F_{j_1}, F_{j_2}) \\ Q \leq F_{j_3} \end{cases} \Leftrightarrow \begin{cases} Q \leq F_{j_1} \\ Q \leq F_{j_2} \\ Q \leq F_{j_3} \end{cases}$$

The same result can be obtained by adopting configuration B. Thus, the optimal value of the flow line scheduling problem with three tasks is the same as the optimal value of the MaxMin problem with three tasks.

For clarity we just index tasks $j, j \leq J + 1$. Suppose now that the optimal value of the flow line with J tasks is the same as the optimal value of the *MaxMin* problem with J tasks. The objective is to show that the flow line with $J + 1$ tasks has the same throughput as the MaxMin problem with the same number of tasks.

Consider now $J + 1$ tasks, the flow line can be described as in Figure J.1:

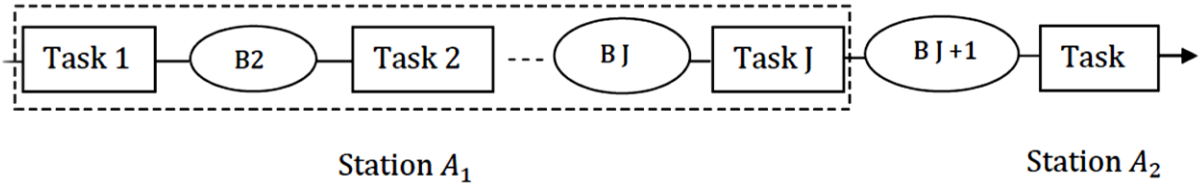


Fig J. 1. *J+1* tasks flow line

Using the same reasoning, solving the flow line with station A_1 and station A_2 is equivalent to solving the *MaxMin* problem with the following constraints:

$$\begin{cases} Q \leq F_{A_1} \\ Q \leq F_{A_2} \end{cases}$$

Using the assumption on the station A_1 and station A_2, we have that:

$$\begin{cases} F_{A_1} \leq F_1 \\ F_{A_1} \leq F_2 \\ \quad \dots \\ F_{A_1} \leq F_J \end{cases} \text{ and } F_{A_2} = F_{J+1}.$$

Inserting the set of constraints on F_{A_1} and F_{A_2} leads to the maximization of Q under the following set of constraints:

$$\begin{cases} Q \leq F_1 \\ Q \leq F_2 \\ \quad \dots \\ Q \leq F_J \\ Q \leq F_{J+1} \end{cases}$$

which defines the *MaxMin* formulation for *J + 1* tasks.

Appendix K. Mixed Integer Nonlinear Program for establishing the upper bound

In this section, we present the MINLP formulation that establishes an upper bound to which we evaluate the performance of the implemented heuristics. We denote the formulation for the parallel system Model 1 and the one for the serial system Model 2.

Model 1: Parallel systems

j task index $j = 1,2, \dots, 9$

i worker index $i = 1,2, \dots, 45$

t index of production periods $t = 1,2, \dots, T$

L period duration

y_{qj} equals 1 if task q and task j are of the same type and 0 otherwise (q is an alias of the task index)

Binary decision Variables

x_{ijt} equals 1 if worker i is assigned to task j at time t and 0 otherwise.

Non negative Variables

S_{jt} the output from task j at period t.

P_{ijt} the productivity rate of worker i assigned to task j at period t .

O_{ijt} the output of worker i on task j at time t.

$$Maximize \sum_{j}^{9} \sum_{t}^{T} S_{jt}$$

$$S_{jt} = \sum_{i}^{45} O_{ijt} \qquad j = 1..9 \qquad t = 1..T \qquad (K-1)$$

$$O_{ijt} = \sum_{k \le t} x_{ijt} * P_{ijk} \qquad i = 1..45 \qquad j = 1..9 \qquad t = 1..T \qquad (K-2)$$

$$T_{ijt} = \sum_{q=1, q \neq j}^{9} y_{qj} * \left(\sum_{p=1, p \neq i}^{9} \left(\sum_{k \le t} x_{pqk} * P_{pqk} \right) \right) \qquad (K-3)$$

$$P_{ijt} = k_{ij} * \frac{\left[\theta_{ij} * T_{ijt} + \left(\sum_{k \le t} L * x_{ijt}\right) + p_{ij}\right]}{\left[\theta_{ij} * T_{ijt} + \left(\sum_{k \le t} L * x_{ijt}\right) + p_{ij} + r_{ij}\right]} \qquad (K-4)$$

$$\sum_{i=1}^{45} x_{ijt} = 1 \qquad j = 1..9 \quad t = 1..T \qquad (K-5)$$

$$\sum_{j=1}^{9} x_{ijt} \le 1 \qquad i = 1..45 \quad t = 1..T \qquad (K-6)$$

Within the parallel production scenario, the objective is to maximize the output from all four tasks. Constraint $(K-1)$ defines the throughput at each task *j* and constraint $(K-2)$ the units produced by each work at each task. The term T_{ijt} that quantifies the knowledge that can be transferred to a worker from others workers when performing similar tasks, is defined in equation $(K-3)$. In equation $(K-4)$ we define the worker productivity rate for a given task at a given period as a function of the transfer parameter, the learning parameters and the cumulate assignments. Constraint $(K-5)$ to constraint $(K-6)$ ensure that at each time *t*, a task is occupied by exactly one worker and a worker is assigned to exactly task.

Model 2: Serial systems

In this second scenario, we consider a flow line with three tasks of different types as described in figure 2. On each of the three flow lines, a unit must progress through the three different tasks sequentially, and this sequence is constant for all jobs. Thus, the objective is to maximize the number of products that are completed at the final task-station. Bentefouet and Nembhard (2012b) establish that the problem, in term of the value of the objective function, is equivalent to the scheduling problem that ignores the buffers constraint and maximizes the least productive tasks. That is, in addition with constraints $(K-1)$ to $(K-6)$, we replace the objective function by:

$$Maximize \;\; Line_1 + Line_2 + Line_3$$

and add the following equations that define the throughput at each of the three lines:

$$Line_1 = \min\left(\sum_{t=1}^{T} S_{1t}, \sum_{t=1}^{T} S_{4t}, \sum_{t=1}^{T} S_{7t}\right) \qquad (K-7)$$

$$Line_2 = \min(\sum_{t=1}^{T} S_{2t}, \sum_{t=1}^{T} S_{5t}, \sum_{t=1}^{T} S_{8t}) \qquad (K-8)$$

$$Line_3 = \min(\sum_{t=1}^{T} S_{3t}, \sum_{t=1}^{T} S_{6t}, \sum_{t=1}^{T} S_{9t}) \qquad (K-9)$$

Appendix L. Top 30 best policies Parallel and Serial system

Table L. 1. Top 30 policy parallel system

Selection	Grouping	Assignment	Mean Throughput	Mean relative gap to the upper bound in % (std in %)
Maximax *k*	Min variance *k*	Maximax *k*	808.89 (3.681)	1.15 (1.03)
Maximax *k*	Min variance *p*	Maximax *Θ*	808.89 (3.681)	1.15 (1.03)
Maximax *k*	Min variance *k*	Maximax *Θ*	807.77 (3.602)	1.15 (1.06)
Maximax *k*	Min variance *1/r*	Maximax *k*	807.77 (3.602)	1.15 (1.06)
Maximax *k*	Min variance *1/r*	Maximax *Θ*	807.77 (3.602)	1.15 (1.06)
Maximax *k*	Min variance *p*	Maximax *k*	806.76 (4.124)	1.16 (1.06)
Maximax *k*	Min variance *Θ*	Maximax *Θ*	806.56 (3.818)	1.16 (0.95)
Maximax *k*	Min variance *Θ*	Maximax *k*	804.88 (6.148)	1.18 (1.03)
Maximax *k*	Min variance *O*	Maximax *p*	803.02 (14.025)	1.54 (1. 36)
Maximax *Θ*	Min variance *O*	Maximax *p*	803.00 (13.715)	1.56 (1.43)
Maximax *Θ*	Min variance *1/r*	Maximax *k*	802.95 (12.356)	1.58 (1.44)
Maximax *k*	Max variance *1/r*	Maximax *Θ*	802.88 (3.773)	1.64 (1.34)
Maximax *Θ*	Min variance *k*	Maximax *Θ*	802.42 (13.094)	1.66 (1.41)
Maximax *Θ*	Min variance *p*	Maximax *Θ*	802.42 (13.094)	1.66 (1.41)
Maximax *Θ*	Min variance *1/r*	Maximax *Θ*	802.42 (13.094)	1.66 (1.41)
Maximax *Θ*	Min variance *Θ*	Maximax *Θ*	802.42 (13.094)	1.66 (1.41)
Maximax *Θ*	Min variance *p*	Maximax *k*	802.25 (13.524)	1.69 (1.44)
Maximax *Θ*	Min variance *Θ*	Maximax *k*	802.25 (13.524)	1.69 (1.44)
Maximax *Θ*	Min variance *k*	Maximax *k*	801.73 (13.264)	1.75 (1.55)
Maximax *Θ*	Min variance *1/r*	Maximax *O*	800.41 (14.652)	1.82 (1.47)
Maximax *Θ*	Min variance *O*	Maximax *Θ*	800.28 (15.569)	1.83 (1.44)
Maximax *k*	Min variance *O*	Maximax *Θ*	800.22 (15.571)	1.84 (1.44)
Maximax *k*	Min variance *k*	Maximax *O*	799.89 (15.810)	1.88 (1.45)
Maximax *k*	Min variance *p*	Maximax *O*	799.89 (15.810)	1.88 (1.45)
Maximax *k*	Min variance *O*	Maximax *O*	799.84 (15.811)	1.89 (1.45)
Maximax *k*	Min variance 1/r	Maximax *O*	799.65 (15.849)	1.91 (1.48)
Maximax *k*	Min variance *Θ*	Maximax *O*	799.65 (15.849)	1.91 (1.48)
Maximax *k*	Min variance *O*	Maximax *k*	799.62 (15.850)	1.91 (1.47)
Maximax *Θ*	Min variance *k*	Maximax *O*	799.20 (15.495)	1.96 (1.57)
Maximax *Θ*	Min variance *O*	Maximax *O*	799.18 (15.495)	1.97 (1.57)

Table L. 2. Top 30 policy serial system

Selection	Grouping	Assignment	Mean Throughput	Mean relative gap to the upper bound in % (std in %)
Maximax Θ	Min variance *1/r*	Maximax *O*	259.07 (7.670)	6.22 (3.14)
Maximax Θ	Min variance *k*	Maximax Θ	258.02 (7.691)	6.28 (3.14)
Maximax Θ	Min variance *k*	Maximax *k*	258.63 (7.812)	6.38 (3.29)
Maximax *k*	Min variance *O*	Maximax *O*	258.62 (7.409)	6.38 (3.03)
Maximax Θ	Max variance *p*	Maximax *k*	258.56 (6.421)	6.42 (2.43)
Maximax *k*	Min variance *k*	Maximax Θ	258.48 (7.275)	6.44 (3.02)
Maximax Θ	Max variance *p*	Maximax Θ	258.40 (6.437)	6.47 (2.42)
Maximax *k*	Min variance *p*	Maximax Θ	258.11 (7.216)	6.57 (2.94)
Maximax *O*	Min variance *k*	Maximax *k*	257.78 (6.994)	6.69 (2. 98)
Maximax Θ	Min variance *O*	Maximax *1/r*	257.63 (12.522)	6.75 (4.73)
Maximax *O*	Min variance *p*	Maximax Θ	257.52 (7.333)	6.78 (3.07)
Maximax *k*	Min variance *k*	Maximax *k*	257.49 (9.257)	6.78 (3.70)
Maximax *O*	Min variance *k*	Maximax *O*	257.29 (7.381)	6.86 (3.07)
Maximax Θ	Min variance *1/r*	Maximax *k*	275.19 (12.74)	6.90 (4.79)
Maximax *O*	Min variance Θ	Maximax *O*	257.11 (7.226)	6.92 (3.01)
Maximax Θ	Min variance *k*	Maximax *O*	257.08 (12.851)	6.94 (4.82)
Maximax Θ	Min variance *O*	Maximax *k*	257.02 (12.851)	6.96 (4.82)
Maximax Θ	Min variance *O*	Maximax Θ	257.01 (12.853)	6.97 (4.82)
Maximax Θ	Min variance *O*	Maximax *O*	256.73 (12.914)	7.07 (4.83)
Maximax *O*	Min variance *p*	Maximax *O*	256.70 (6.903)	7.08 (2.89)
Maximax *O*	Min variance Θ	Maximax *p*	256.34 (6.764)	7.20 (3.01)
Maximax Θ	Max variance *1/r*	Maximax *k*	256.15 (6.066)	7.28 (2.41)
Maximax Θ	Min variance Θ	Maximax *k*	256.07 (12.693)	7.30 (4.86)
Maximax Θ	Min variance *p*	Maximax Θ	256.04 (12.399)	7.31 (4.75)
Maximax Θ	Min variance *p*	Maximax *k*	256.01 (12.135)	7.32 (4.66)
Maximax *k*	Min variance *p*	Maximax *k*	256.00 (12.551)	7.32 (4.68)
Maximax *k*	Min variance *O*	Maximax *k*	255.97 (12.421)	7.34 (4.63)
Maximax *k*	Min variance *1/r*	Maximax *k*	255.47 (16.875)	7.52 (6.20)
Maximax *k*	Min variance *k*	Maximax *O*	255.28 (13.815)	7.60 (5.10)
Maximax *k*	Min variance *1/r*	Maximax Θ	255.22 (17.231)	7.62 (6.33)

Bibliography

Aardal, K., and Ari, A. 1987. Decomposition Principles Applied to the Dynamic Production and Workforce Scheduling Problem. Engineering Cost and Production Economics, 12(1-4), 39-49.

Abernathy, W.J., Clark, K.B., and Kantorow, A.M. 1981. The new industrial competition. Hav. Bus. Rev., 59, 68-81.

Aggarwal, S.C. 1982. A focused review of scheduling in services. European Journal of Operational Research 9(2), 114-121.

Alvarez-Valdes, Crespo, R., E., and Tamarit, J. 1999. Labour Scheduling at an Airport Refuelling Installation. Journal of the Operational Research Society 50(3), 211–218.

Anzanello, M.J., Fogliatto, F.S. 2007. Learning curve modeling of work assignment in mass customized assembly lines. International Journal of Production Research 45 (13), 2919–2938.

Anzai, Y. and H. A. Simon. 1979. The theory of learning by doing. Psychological Review 86:124-140.

Argote, L. 1999. Organizational learning:Creating, retaining and transferring knowledge. Norwell, MA: Kluwer.

Armbruster, D., Gel, E.S., Murakami, J., 2007. Bucket brigades with worker learning. European Journal of Operational Research 176 (1), 264–274.

Arthur, J. and Ravindran, A. 1981. "A Multiple Objective Nurse Scheduling Model." AIIE Transactions 13(1), 55–60.

Arzi, Y. and Shtub, A. 1997. Learning and Forgetting in mental and mechanical tasks: a comparative study. IIE Transactions 29(9),759-768.

Atlason, J., Epelman, M., and Henderson, S. 2002. Combining Simulation and Cutting Plane Methods in Service Systems. In Proceedings of the 2002 National Science Foundation Design, Service and Manufacturing Grantees Conference.

Atlason, J., Epelman, M.A., and Henderson, S.G. 2004. Call Center Staffing with Simulation and Cutting Plane Methods. Annals of Operations Research 127, Special Issue on Staff Scheduling and Rostering, 333–358.

Avramidis, A.N., Chan, W., and L'Ecuyer, P. 2009. Staffing multiskill call centers via search methods and a performance approximation. IIE Transactions 4 (1), 483-497.

Badiru, A.B. 1992. Computational survey of univariate and multivariate learning curve models. IEEE Trans. Eng. Manage. 39 (2), 176–188.

Badiru, A.B., Multifactor learning and forgetting models for productivity and performance analysis, Int.J. Hum. Factor. Man., 4, 37–54, 1994.

Baloff, N. 1966. Learning Curves- Some Controversial Issues. Journal of Industrial Economics. 14, 275-282.

Baloff, N. 1971. Extension of the Learning Curve-Some Empirical Results. Operation Research Quaterly, 22,329-340.

Bailey, C.D.1989. Forgetting and the learning curve: a laboratory study, *Manage. Sci.*, 35, 340–352.

Baines, T.S., Kay, J.M., 2002. Human performance modeling as an aid in the process of manufacturing system design: a pilot study. International Journal of Production Research 40 (10), 2321–2334.

Baloff, N. 1970. Startup Management. IEEE Trans Engineering Management, EM-17, 132-141.

Baker, K. 1974. Introduction to Sequencing and Scheduling. Wiley. New York.

Baker, K., and Magazine, M. 1977. Workforce Scheduling with Cyclic Demands and Day-off Constraints. Management Science 24(2), 161–167.

Baker, K., Powell, S., Pyke, D., 1990. Buffered and unbuffered assembly systems with variable processing times. Journal of Manufacturing and Operation Management 3, 200–223.

Barachini, F. 1994. Frontiers in run-time prediction for the production system paradigm. AI Mag. 15 (3), 47-61.

Bartholdi, J.J., Bunimovich, L.A., Eisenstein, D.D., 1999. Dynamics of 2- and 3- worker bucket brigade production lines. Operations Research 47, 488–491.

Barichard, V., Ehrgott, M., Gandibleux, X. and T'Kindt, V. 2009. Multiobjective Programming and Goal Programming. Theoretical Results and Practical Applications. Serie: Lecture Notes in Economics and Mathematical Systems, Vol. 618. Springer.

Bartholdi III, J.J., Eisenstein, D.D., 1996. A production line that balances itself. Operational Research 44 (1), 21–34.

Bartholdi, J.J., Eisenstein, D.D., Foley, R.D., 2001. Performance of bucket brigades when work is stochastic. Operations Research 49 (5), 710–719.

Baum, J.A.C. and Ingram, P. 1998. Survival-enhancing learning in the Manhattan hotel industry 1898-1980. Management Science, 44 (7), 996-1016.

Bechtold, S., Brusco, M., and Showalter, M. 1991. A Comparative Evaluation of Labor Tour Scheduling Methods. Decision Science, 22, 4. September-October 1991.

Bechtold, S., Janaro, R., and Sumners, D.W. 1984. Maximization of Labor Productivity Through Optimal Rest-Break Schedules. Management Science, 30(12), 1442-1458.

Becker,M., Grabot,B.,Laring,J.,Mummolo,G..,Zu€ lch , G.,2004.Human-Centered Simulation: AnalysisofStateoftheArt,ProspectsandBenefits.Sim-Serv, Espoo //http://www.sim-serv.com/white_papers. phpS.

Bellman, R.E., 1957. Dynamic Programming. Princeton, NJ: Princeton University Press.

Bentefouet, F. and Nembhard, D.A, 2012b. Optimal Flow-Line Conditions with Worker Variability. International Journal of Production Economics. http://dx.doi.org/10.1016/j.ijpe.2012.10.008.

Bhaskar, K. and Srinivasan, G. 1997. Static and dynamic operator allocation problems in cellular manufacturing systems. International Journal of production Research 35 (12), 3467-3481.

Bhulai, S., Koole, G. and Pot, A. 2008. Simple Methods for Shift Scheduling in Multiskill Call Centers. Manufacturing & Service Operations Management 10(3), 411-420.

Biskup, D. 1999. Single-machine scheduling with learning considerations, Eur. J. Oper. Res., 115, 173–178.

Biskup, D. 2008. A state-of-the-art review on scheduling with learning effects. European Journal of Operation Research 188, 315-329.

Borst, S. 2001. Robust Algorithms for Sharing Agents with Multiple Skills. Technical Report, Bell Laboratories.

Boudreau, J., Hopp, W., McClain, J.O., and Thomas, L.J. 2003. On the interface between operations and human resources management. Manufacturing & Service Operations Management, 5 (3), 179-202.

Bratcu, A. and Dolgui, A. 2005. A survey of the self-balancing production lines: Bucket Brigades, Journal of Intelligent Manufacturing 16 (2), 139-158.

Browne, S. and Yechiali, U. 1990. Scheduling Detteriorating Jobs on a Single Processor. Operations Research, 38(3), 495-498.

Burke, E., De Causmaecker, P., Vanden Berghe, G., and Van Landeghem, H. 2002. The State of the Art of Nurse Rostering. Technical Report, School of Computer Science and IT, University of Nottingham.

Buxey, G. 1989. Production scheduling: Theory and practice. European Journal of Operational Research, 39(1), 17–31.

Buzacot, J.A. 2002. The impact of worker differences on production systems output. International Journal of Production Economics 78, 37-44.

Campbell, G.M. 1999. Cross-utilization of workers whose capabilities differ. Management Science, 45(5), 722-732.

Campbell, G. 2002. Development and Evaluation of an Assignment Heuristic for Allocating Cross-Trained Workers. European Journal of Operational Research 138, 9–20.

Carlson, J. G. 1961. How Management Can Use the Improvement Phenomenon. California Management Review, 3(2), 83-94.

Carlson, J.G. and Rowe, R.J.1976. How much does forgetting cost? Ind. Eng., 8, 40–47.

Carraway, R., 1989. A dynamic programming approach to stochastic assembly line balancing. Management Science 35 (4), 459–471.

Carraway, R.L., Morin, T.L., and Moskowitz, H. 1990. Generalized dynamic programming for multicriteria optimization. European Journal of Operational Research 44, 95-104.

Cezik, M. T., L'Ecuyer, P. 2008. Staffing multi-skill call centers via linear programming and simulation. Management Sci. 54(2), 310-323.

Chen, J. and Yeung, T. 1992. "Development of a Hybrid Expert System for Nurse Shift Scheduling. International Journal of Industrial Ergonomics 9(4), 315–327.

Chen, T.C.E., Wu, C.C., and Lee, W.-C. 2008. Some scheduling problems with sum-of-processing – Times-based and job-position-based learning effects. Information sciences 178(11), 2476-2487.

Cheng, T.C.E., Wu, C.-C., Chen, J.-C., Wu, W.-H., Cheng, S.-R., 2012. Two-machine flowshop scheduling with a truncated learning function to minimize the makespan. International Journal of Production Economics http://dx.doi.org/ 10.1016/j.ijpe.2012.03.027.

Cochran, E.B. 1960. New concepts of the learning curve. The Journal of Industrial Engineering 11, 317-327.

Conway, R.W. and Schultz, A. 1959. The manufacturing progress function. The journal of Industrial Engineering 10, 39-54.

Corominas, A., Olivella, J., and Pastor, R. 2010. A model for the assignment of a set of tasks when work performance depends on experience of all tasks involved. International Journal of production Economics 126, 335-340.

Corominas, A., Pastor, R., and Rodriguez, E. 2006. Rotational allocation of tasks to multifunctional workers in a service industry. International Journal of production Economics 103, 3-9.

Cox, T., Griffiths, A. 1994. The nature and measurement of work stress: theory and practice. In: Wilson, J., Corlett, E. (Eds.). Evaluation of Human Work: a Practical Ergonomics Methodology. CRC Press, pp. 783‑803.

Crawford, S., and Wiers, V.C.S. 2001. From anecdotes to theory: a review of existing knowledge on human factors of planning and scheduling. In MacCarthy, B.L. & Wilson, J.R. (Eds.), Human performance in planning and scheduling, (pp. 15-43). London: Taylor and Francis.

Crossman, E.R.F.W. 1959. A theory of acquisition of speed skill, Ergonomics, 2, 153–166.

Czeisler, C.A., Moore-Ede, M.C., Coleman, R.C. 1982. Rotating shift work schedules that disrupt sleep are improved by applying circadian principles. Science 217, 460‑463.

Dar-El EM (2000). Human Learning: From Learning Curves to Learning Organizations. Kluwer: Boston, MA.

Dar-El, E. M., Ayas, K., and Gilad, I. 1995. A Dual-Phrase Model for the individual learning process in industrial learning process in individual tasks. IIE Trans. 27(3), 265–271.

Darr, E., Argote, L., and Epple, D. 1995. The acquisition, transfer, and depreciation of learning in service organizations: Productivity in franchises. Management Science, 44, 1750-1762.

Dejong, J.R. 1957. The effects of increasing skill on cycle time and its consequence for time standards. Ergonomics 51-60.

Dessouky, M.I., Moray, N., Kijowski, B.A. 1995. Taxonomy of scheduling systems as a basis for the study of strategic behavior. Human Factors 37 (3), 443‑472.

Doerr, K.H.,Arreola-Risa,A.,2000.Aworker-basedapproachformodeling variability in task completion time. IIE Transactions 32, 625–636.

Doerr, K.H.,Freed,T.,Mitchell,T.R,Schriesheim,C.A.,Zhou,X.,2004.Work flow policy and within-worker and between- workers variability in performance. Journal of Applied Psychology 89(5), 911–921.

Dudek, R. A., Panwalker, S. S., and Smith, M. L. 1992. The lessons of flow shop scheduling Research. Operations Research, 40(1), pp. 7–13.

Dutton, J.M. 1962. Simulation of an actual production scheduling and workflow control system. International Journal of Production Research, 4 ,21-41.

Ebbinghaus, H. 1885. A study in logical memory. Memory (Trans. by Ruger and Bussenius, 1913).

Ebeling, A.C., and Lee, C.Y. 1994. Cross-training Effectiveness and Profitability. International Journal of Production Research 32 (12), 2843-2859.

Edie, L. 1954. Traffic Delays at Toll Booths. Journal Operations Research Society of America 2(2), 107–138.

Ernst, A.T., Jiang, H., Krishnamoorthy, M., Owens, B., and Sier, D. 2004. An Annotated Bibliography of Personnel Scheduling and Rostering. Annals of Operations Research 127(1-4), 21-144.

Eskildsen, J.K., Kristensen, K., Westlund, A.H. 2004. Work motivation and job satisfaction in the Nordic countries. Employee Relations 26 (2), 122‑136.

Faraj, S. and Sproull, L. 2000. Coordinating expertise in software development teams. Manage. Sci., 46, 1554-1568.

Fine, C.H. 1986. Quality improvement and learning in productive systems. Manage. Sci., 32, 1302-1315.

Finkelman, J. 1994. A large database study of the factors associated with work-induced fatigue Human Factors, 36(2), 232-243

Fox, M. S., and Smith, S. F. 1984. ISIS: A knowledge based system for factory scheduling. Expert Systems, 1,25-49.

Fry, T.D., Kher, H.V., and Malhotra, M.K. 1995. Managing worker flexibility and attrition in dual resource constrained job shops. International Journal of Production Research, 33(8), 2163-2179.

Fudenberg, D. and Tirole, J. 1983. Learning by doing and market performance. Bell Journal of Economics 14(2), 522-530.

Gagnon, R., Ghosh, S., 1991. Assembly line research: historical roots, research life cycles and future directions. OMEGA International Journal of Management Science 19, 381–399.

Gerwin, D. 1993. Manufacturing Flexibility: A Strategic Perspective. Management Science 39(4), 395-410.

Gherardi, S. and Nicolini, D. The sociological foundations of organizational learning. In M. Dierkes, A.B. Antal, J.Child and I. Nomaka (Eds), Handbook of organizational learning and knowledge. Oxford: oxford University Press, 2001, pp.35-60.

Glazebrook, K.D. and Mitchell, H.M. 2002. An index policy for a stochastic scheduling model with improving/deteriorating jobs. Naval Research Logistics, 49(7), 706-721.

Globerson, S. 1980. The influence of job related variables on the predictability power of three learning curve models. AIIE Trans.12(1), 64–69.

Globerson, S., Levin, N., Shtub, A. 1989. The impact of breaks on forgetting when performing a repetitive task. IIE Transactions 21 (4), 376}381.

Globerson, S., Nahumi, A., and Ellis, S. 1998. Rate of forgetting for motor and cognitive tasks, International Journal Cognitive Ergonomics, 2, 181–191.

Graves, S. C. 1981. A review of production scheduling. Operations Research, 29(2), pp. 646–675.

Gupta, J.N.D. and Gupta, S.K. 1988. Single facility scheduling with nonlinear processing times. Computers & Industrial Engineering, 14(4), 387-393.

Hancock, W.M. 1967. The prediction of learning rates for manual operations. Journal of Industrial Engineering 18 (1), 42-47.

Hao, G., Lai, K.K., and Tan, M. 2004. A Neural Network Application in Personnel Scheduling. Annals of Operations Research 128, Special Issue on Staff Scheduling and Rostering, 65–90.

Hertzberg, F., Mausner, B., Snyderman, B.1959. The Motivation to Work, second ed. Wiley.

Hopp, W. J., and Van Oyen, M.P. 2004. Agile workforce evaluation: aA framework for cross-training and coordination. IIE Trans. 36, 919-940.

Huang, X.,Wang, J.-B., Wang, L.-Y, Gao, W.-J and Wang, X.-R. 2010. Single machine scheduling with time-dependent deterioration and exponential learning effect. Computers & Industrial Engineering 58, 58-63.

Hunter, J.E., Schmidt, F.L., Judiesch, M.K., 1990. Individual differences in output variability as a function of job complexity. Journal of Applied Psychology 75, 28–42.

Hurst, E. G., and McNamara, A. B. 1967. Heuristic scheduling in a woolen mill. Management Science, 14,182-203.

Ingolfsson, A., Cabral, E. and Wu, X. 2007. Combining integer programming and the randomization method to schedule employees. Technical report, School of Business, University of Alberta, Edmonton, Alberta, Canada.

Kanet, J. J., and Adelsberger, H. H. 1987. Expert systems in production scheduling. European Journal of Operations Research, 29, 5 I-59.

Kapp, K.M. 1999. Transforming your manufacturing organization into a learning organization, Hosp. Mater. Manage. Q., 20, 46-54.

Kavinoky,R. and Thoo,J.B., Spring 2008. The number of real roots of a cubic Equation. The AMATYC Review 29 (2), pp. 3–8.

Keachie, E.C. and Fontana, R.J. 1966. Effect of Learning on Optimal Size. Management Science, B102-108.

Kim, D.H. 1993. The link between individual and organizational learning. Sloan Management Review, Fall, 37-50.

Kim, S. and Nembhard, D.A. 2010. Cross-trained staffing levels with heterogeneous learning/forgetting. IEEE Transactions on Engineering Management 57 (4), 560-574.

Konz, S. 1983. Work Design: Industrial Ergonomics. 2nd eds, New York, Wiley.

Kostreva, M. and Jennings, K. 1991. Nurse Scheduling on a Microcomputer. Computers and Operations Research 18(8), 731–739.

Kostreva, M., McNelis, E. and Clemens, E. 2002. Using a circadian rhythms model to evaluate shift schedules, Ergonomics 45(11), 739–763.

Kottas, J., Lau, H., 1981. A stochastic line balancing procedure. International Journal of Production Research 19, 177–193.

Koulamas, C. and Kyparisis, G.J. 2007. Single-machine and two machine flowshop scheduling with general learning function. European Journal of Operational research 187, 68-72.

Knauth, P. 1993. The design of shift systems. Ergonomics 36, 15‑28.

Lac, G., Chamoux, A. 2004. Biological and psychological responses to two rapid shiftwork schedules. Ergonomics 47 (12), 1339‑1349.

Larson, J.R., Christensen, C., Abbott, A.S., and Franz, T.M. 1996. Diagnosing groups: Charting the flow of information in medical decision-making teams. Journal Personality Social Psychology, 71, 315-330.

Leung, J.Y-T. 2004. Handbook of Scheduling: Algorithms, Models and Performance Analysis. Chapman & Hall/CRC, Boca Raton, Fl, 1120 pp.

Levy, F. K. 1965. Adaptation in the production process. Management Science,11, B136-B154.

Levitt, B., and March, J.G. 1988. Organizational learning. Annual Review of Sociology, 14, 319-340.

Lewis, K. 2003. Measuring transactive memory systems in the field: Scale development and validation. Journal Applied Psychology, 88(4), 587-604.

Liang, D. W., Moreland, R., and Argote, L. 1995. Group versus individual training and group performance: The mediating role of transactive memory. Personality Social Psychology Bulletin, 21, 384-393

Lieberman, M.B. 1984. The Learning curve and pricing in the chemical processing industries. Rand Journal of Economics, 15(2), 213-228.

MacCarthy, B. L. and Liu, J. 1993. Addressing the gap in scheduling research: A review of optimization and heuristic methods in production scheduling. International Journal of Production Research, 31(1), pp. 59–79.

Malhotra, M.K., Fry, T.D., Kher, H.V., and Donohue, J.M. 1993. The impact of learning and labor attrition of worker flexibility in dual resource constrained job shops. Decision Sciences, 24(3), 641-663.

Malladi, S. and Jo Min, K. 2004. Workforce scheduling with costs and ergonomic considerations. Proceedings of the Industrial Engineering Research Conference, CD-ROM. Atlanta, GA.

Mandelbaum, A. and Zeltyn, S. 2006. Service engineering of call centers: Research, teaching, practice. IBM Business Optimization and Operation Research Workshop, Haifa, Israel.

March, J.G. 1991. Exploration and exploitation in organizational learning. Orgnizational Science, 2(1), 71-87.

Mazur, J.E. and Hastie, R. 1978. Learning as Accumulation: A Reexamination of the Learning Curve. Psychological Bulletin 85 (6), 1256-1274.

Mazzola, J.B. and Neebe, A.W. 1988. Bottleneck Generalized Assignment Problems. Engineering Cost and Production Economics 14 (1), 61-66.

McDonald, T., Ellis, K., Van Aken, E., and Patrick Koelling, C. 2009. Development and application of a worker assignment model to evaluate a lean manufacturing cell. International Journal of production Research 47 (9), 2427-2447.

McKay,K.N, Safayeni, F.R., and Buzacott, J.A. 1988. Job-shop scheduling theory: what is relevant? Interface, 18, 84-90.

McKay,K.N, Buzacott, J.A., and Safayeni, F.R. 1989. The scheduler's knowledge of uncertainty: the missing link in J. Browne 9ed.) Knowledge Based Production Management Systems, IFIP (Elsevier Science Publishers B. V., North Holland).

McKay,K.N.,Morton,T.E. and RammathP. 1995. Aversion dynamics: scheduling when the system changes. Working Paper, University of Waterloo.

McKay, K.N., and Wiers, V.C.S. 1999. Unifying the theory and practice of production scheduling. Journal of Manufacturing Systems 18(4), 241-255.

Mitten, L.G. 1970. Branch and bound methods: general formulation and properties. Operations Research 18, 24-34.

Mehrotra, V., Fama, J. 2003. Call centers simulation modeling: Methods, challenges, and opportunities in Proceedings of the 35th conference on Winter Simulation, New Orleans, LA, S. Chick.

Meilijson, I., and Tamir, A. 1984. Minimizing Flow Time On Parallel Indetical Processors with Variable Unit Processing Time. Opns. Res. 32, 440-448.

Moodie, C., Young, H., 1965. A heuristic method of assembly line balancing for assumptions of constant or variable work element times. Journal of Industrial Engineering 16 (1), 23–29.

Molleman, E. and Slomp, J. 1999. Functional flexibility and team performance. International Journal of Production Research 37, 1837-1858.

Monk, T.H. 2000. What can the chronobiologist do to help the shift worker? Journal of Biological Rhythms 15, 86‑94.

Moray, N., Dessouky, M.I., Kijowski, B.A., Adapathya, R. 1991. Strategic behavior, workload, and performance in task scheduling. Human Factors 33 (5), 607‑629.

Moreland, R.L., Argote, L., and Krishnan, R., 1996. Socially shared cognition at work: Transactive memory and group performance. J.L. Nye, A.M. Brower, eds. *What's So Social About Cognition? Social Cognition Research in Small Groups*. Sage, thousand Oaks, CA, 57-84.

Mosheiov, G. and Sidney, J.B. 2003. Note on scheduling with general learning curves to minimize number of tardy jobs. Journal of the Operational Research Society 56, 110-112.

Nembhard, D.A. 2001. Heuristic approach for assigning workers to task based on individual learning rates. International Journal of Production Research 39 (8), 1968-1995.

Nembhard, D.A, Bentefouet, F., 2012a. Parallel system scheduling with general learning and forgetting. International Journal of Production Economics.

Nembhard, D.A and Norman, B.A. 2007. Cross training in production systems with human learning and forgetting. In: Workforce Cross-Training. CRC Press, Florida, 111-129.

Nembhard, D.A. and Osothsilp, N. 2001. An empirical comparison of forgetting models. IEEE Trans. Eng. Manage. 48 (3), 283-291.

Nembhard, D.A., Osothslip, N., 2005. Learning and forgetting-based worker selection for tasks of varying complexity. Journal of the Operational Research Society 56, 576–587.

Nembhard, D.A and Shafer, S.M. 2007. The effects of workforce heterogeneity on productivity in an experiential learning environment. The International Journal of Production Research 46 (14), 3909-3929.

Nembhard, D.A. and Uzumeri, M.V. 2000. An individual based description of learning within an organization. IEEE Trans. Eng. Manage. 47 (3), 370-378.

Nembhard, D.A. and Uzumeri, M.V. 2000a. Experiential Learning and Forgetting for Manual and Cognitive Tasks. International Journal of Industrial Ergonomics 25 (4), 370-378.

Nembhard, D.A. and Uzumeri, M.V. 2000b. Experiential Learning and Forgetting for Manual and Cognitive Tasks. International Journal of Industrial Ergonomics 25 (4), 315-326.

Neves, D.M., Anderson, J.R. 1981. Knowledge compilation: mechanisms for the automatization of cognitive skills. In: Anderson, J.R. (Ed.), Cognitive Skills and Their Acquisition. Erlbaum, Hillsdale, NJ, USA.

Newell, A. and Rosenbloom, P.S. 1981. Mechanisms of skills acquisition and the law of practice. In J.R. Anderson (Ed.),Cognitive skills and their acquisition. Hillsdale, NJ: Erlbaum.

Norman, B.A., Tharmmaphornphilas, W., Needy, K.L., Bidanda, B., and Warner, R.C. 2002. Worker assignment in cellular manufacturing considering technical and human skills. International Journal of Production Research 40 (6), 1479-1492.

Olivella, J. 2007. An experiment on task performance forecasting based on the experience of different tasks. I-KNOW'07, Graz 305-312.

Park, P.S, and Bobrowski, P.M. 1989. Job release and labor flexibility in a Dual Resource Constrained job shop. Journal of Operation Management, 88, 230-249.

Pinedo, M. 1995. Scheduling: Theory, Algorithms, and Systems, 1st edition. Prentice-Hall.

Pinheiro, J.C., and Bates, D.M. 2000 . Mixed Effects Models in S and S-Plus. Spring-Verlag: New York Inc., NY.

Powell, S. 1992. Buffer Allocation in Unbalanced Serial Lines. Working Paper 289, The Amos Tuck School of Business Administration, Dartmouth College, Han- over, NH, USA.

Pratsini, E., Camm, J.D., and Raturi, A.S. 1993. Effect of Process Learning on Manufacturing Schedules. Computer Ops Res. 20, 15-24.

Qin, R. and Nembhard, D. A. 2007. Valuing workforce cross-training flexibility, In: Workforce Cross-Training. CRC Press, Florida, 131-151.

Ramos, M. A. G. and Chen, J.G. 1994. On the integration of learning and forgetting curves for the control of knowledge and skill acquisition for non-repetitive task training and retraining. Int. J. Ind. Eng. 1(3),233–242.

Rapping, L. 1965. Learning and World War II production functions. Review of Economics and Statistic,48(1), 81-86.

Rockafellar, R.T., 1970. Convex analysis. Princeton: Princeton University Press.

Rodammer, F.A., and White, K.P. 1988. A recent survey of production scheduling. IEEE Transaction on Systems, Man, and Cybernetics, 18(6).

Rosekind,M.R., Graeber, R.C., Dinges, D.F., Connell, L.J., Rountree, M.S., and Gillen, K. 1994. Crew Factors in Flight Operations IX: Effects of Planned Cockpit Rest on Crew Performance and Alertness in Long-Haul Operations (NASA Technical Memorandum 108839). Nasa Ames Research Center, California

Rothe, H.F., 1978. Output rates among industrial employees. Journal of Applied Psychology 63, 40–46.

Ryan, T. and Joiner, B. 1973. Normal Probability Plots and test for normality. Tech. Rept., The Pennsylvania State University, University Park, PA.

Saaty, T.L. 1980. The Analytic Hierarchy Process, New York: McGraw Hill. International,

Translated to Russian, Portuguese, and Chinese, Revised editions, Paperback (1996, 2000), Pittsburgh: RWS Publications.

Sarin, S.C., Nagarajan, B., and Liao, L. 2010. Stochastic scheduling. Cambridge University Press. Cambridge, New York.

Sarin, S.C., Erel, E., 1991. Development of cost model for the single-model stochastic assembly line balancing problem. International Journal of Production Research 28 (2), 1305–1316.

Schmidt, F.L.,Hunter,J.E.,Outerbridge,A.N.,Goff,S.,1988.Joint relation of experience and ability with job performance: test of three hypotheses. Journal of Applied Psychology 73, 46–57.

Shafer, S.M., Nembhard, D.A., and Uzumeri, M.V. 2001. Investigation of Learning Forgetting, and Worker Heterogeneity on Assembly Line Productivity. Management Science 47 (12), 1639-1653.

Sheshinski, E. 1967. Test of the learning by doing hypothesis. Review of Economics and Statistics, 49(4), 568-578.

Shtub, A., Levin, N., and Globerson, S.1993. Learning and forgetting industrial tasks: an experimental model, Int. J. Hum. Factor. Man., 3, 293–305.

Simon, H. 1991. Bounded rationality and organizational learning. Organizational Science, 1(2), 125-134.

Smith-Jentsch, K.A., Zeisig, R.L. Acton, B., and McPherson, J.A. 1998. Team dimensional training: a strategy for guided team self-correction. In: Making Decision under stress: Implications for Individual and Team Training, J.A. Cannon Bowers and E. Salas, eds. Washington, D.C.: American Psychological association, 271-297.

Snoddy, G. S. 1926. Learning and stability. Journal of Applied Psychology, 10, 1–36.

Snook, S.H. 1978. The design of manual handling tasks. Ergonomics, 21, 963-985.

Spence, A.M. 1981. The learning curve and competition. Bell Journal of Economics, 12(1),49-70.

Steward, B.D, Webster, D.B., Ahmad, S., and Matson, J.O. 1994. Mathematical models for developing a flexible workforce. International Journal of Production Economics 36, 243-254.

Stobaugh, R.B. and Townsend, P.L. 1975. Price forecasting and strategies planning: the case of petrochemicals. Journal of Marketing Research, XII, 19-20.

Tepas, D.I., Paley, M., Popkin, S. 1997. Work schedules and sustained performance. In: Salvendy, G. (Ed.), Handbook of Human Factors and Ergonomics, second ed. John Wiley and Sons, New York, Chichester, Brisbane, 1021‑1058.

Teyarachakul, S., Chand, S., and Ward, J. 2011. Effect of learning and forgetting on batch sizes. Production and Operations Management 20 (1), 116-128.

Tharmmaphornphilas, W., and Norman, B.A. 2004. A Quantitative Method for Determining Proper Job Rotation Intervals. Annals of Operations Research 128, Special Issue on Staff Scheduling and Rostering.

Thomas, B.G. and Nembhard, D.A. 2004. Preference based search approach for scheduling workers with learning and forgetting. In: Proc. MSOM Sponsored Session. INFORMS Ann. Meeting, Denver, Oct.

Tietjen, M.A., Myers, R.M. 1998. Motivation and job satisfaction. Management Decision 36 (4), 226‑231.

Tien, J., and Kamiyama, A. 1982. On manpower scheduling algorithms. SIAM Review, *24*(3), 275–287.

Uzumeri, M.V. and Nembhard, D.A. 1998. A population of learners: a new way to measure organizational learning. J. Oper. Manage. 16 (5), 515-528.

Uzzi, B., 1996. Sources and consequences of embeddedness for the economic performance of organizations. American Sociological Review, 61, 674-698.

Uzzi, B.,Lancaster,R.,2003. Relational embeddedness and learning: The case of bank loan managers and their clients. Management science, 49, 383-399.

Van Cott, H.P. and Kinkade, R.G. 1972. Human engineering guide to equipment design (Rev.ed.). Washington D.C.: U.S. Government Printing Office.

Venezia, I. 1985. On the statistical origins of the learning curve. European Journal of Operation Research 19, 191-200.

Volpe, C., Cannon-Bowers, J.A., Salas, E., and Spector, P.E. 1996. The impact of cross-training on team functioning: an empirical investigation. Human Factors 38, 87-100.

Voutsinas, T.G. and Pappis, C.P. 2002. Scheduling jobs with values exponentially deteriorating over time. Internationla Journal of Production Economics, 79(3), 163-169.

Wang, F., Qi, L., 2005. Comments on ‘‘ explicit criterion for the positive definiteness of a general quartic form. IEEE Transactions on Automatic Control 50 (3), 416–418.

Warner, R.C., Needy, K.L., and Bidanda, B. 1997. Worker assignment in implementing manufacturing cells. In: Proceedings of the Sixth Industrial Engineering Research Conference. Miami Beach, FL, 240-245.

Wegner, D.M. 1986. Transactive memory: A contemporary analysis of the group mind. B. Mullen, G. R. Goethals, eds. *Theories of Group behavior*. Springer-Verlag, New-York, 185-205.

Wegner, D.M. 1995. A computer network model of human transactive memory. *Social Cognition*, **13**, 319-339.

Westfall, M., September 1996. Too tired to work? Fatigue presents a major obstacle to safety. Safety & Health, 74‑76.

Wickens, C.D., Lee, J., Liu, Y.D., Gordon-Becker, S. 2004a. Stress and Workload, in Introduction to Human Factors Engineering, second ed. Prentice Hall, Upper Saddle River, New Jersey, USA, pp. 377‑408 (Chapter 13).

Wickens, C.D., Lee, J., Liu, Y.D., Gordon-Becker, S. 2004b. Automation, in Introduction to Human Factors Engineering, second ed. Prentice Hall, Upper

Saddle River, New Jersey, USA, pp. 493‑512 (Chapter 16).Wisner, J.D., and Pearson, J.N. 1993. An exploratory study of the effects of operator relearning in a dual resource constrained job shop Production and Operations Management Journal, 2, 55–68

Wilhelm, W., 1987. On the normality of operation times in small lot assembly systems: a technical note. International Journal of Production Research 25 (1), 145–149.

Wisner, J.D. and Pearson, J.N. 1993. An exploratory study of the effects of operator relearning in a dual resource constrained job shop. Prod Opns Mgmt 2,55-68.

Wright, T. P. 1936. Factors Affecting the Cost of Airplanes. Journal of the Aeronautical Sciences, 3, pp. 122-128.

Wu, C.-C., Wu, W.H., Hsu, P.-H., Lai, K., 2011. A two-machine flowshop scheduling problem with a truncated sum of processing-times- based learning function. Applied Mathematical Modelling http://dx.doi.org/10.1016/ j.apm. 2011.038.

Wu, C.-C., Yin, Y., Cheng, S.-R., 2012. Single-machine and two-machine flowshop scheduling problems with truncated position-based learning functions. Journal of the Operation Research Society, 1–10.

Yelle, L.E. 1979. The Learning Curve: Historical Review and Comprehensive Survey. Decision Science 10 (2), 302-328.

Yin, Y., Wu, C.C., Wu, W.H., Cheng, S.R., 2012. The single-machine total weighted tardiness scheduling problem with learning effects. Computers &Operations Research 39 (5), 1109–1116.

Zangwill, W.I. and Kantor, P.B. 1998. Toward a theory of continuous improvement and learning curves. Manage. Sci, 44, 910-920.

Frank BENTEFOUET

Frank.bentefouet@centraliens.net

EDUCATION

Ph.D. Industrial Engineering and Operation Research, 2009-2013, The Pennsylvania State. University. Advisor: Professor David. A. Nembhard.

Stochastic Calculus and Financial derivatives, Master El Karoui & Ecole Polytechnique, France.

M. Eng, Financial Engineering, 2008, Ecole Centrale Paris, France.

B.S., Applied Math, 2004, University Pierre et Marie Curie Paris VI, France.

B.S., Pure Math, 2002, University Cheikh Anta Diop de Dakar, Senegal.

PUBLICATIONS

Nembhard, D.A. and **Bentefouet, F**.,"Parallel system scheduling with general worker learning and forgetting", *International Journal of Production Economics.*

Bentefouet, F. and Nembhard, D.A., "Optimal Flow-Line conditions with worker variability", *International Journal of Production Economics.*

RESEARCH EXPERIENCE

Research Assistant, 2009-2011, Workforce Engineering Laboratory, Penn State University.

Advisor: Dr. David. A. Nembhard, Department of Industrial and Manufacturing Engineering.

Research Assistant, 2007, Applied Math Laboratory, Ecole Centrale Paris.

Advisor: Dr. Paul-Henry Cournede, Applied Math Laboratory, Ecole Centrale Paris.

PRESENTATIONS

Bentefouet, F. and Nembhard, D.A, "Optimal Flow-Line conditions with worker variability", *INFORMS Conference*, invited session, Phoenix, Arizona. Fall 2012, **presenter**.

Bentefouet, F. and Nembhard, D.A, "Parallel System Scheduling with Learning and Forgetting", *INFORMS Conference*, invited session, Charlotte, North Carolina. Fall 2011. **Presenter.**