An overview of flexibility literature from the operations management perspective

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

**An overview of flexibility literature from
the operations management
perspective**
J.W.M. Bertrand
WP 86

An overview of flexibility literature from the operations management perspective

J.W.M. Bertrand
Subdepartment Operations Planning and Control
Department of Technology Management

An overview of flexibility literature from the operations management perspective

1. Introduction

After costs, quality and reliability, flexibility has emerged in the last two decades as the fourth important performance indicator for operational systems. Researchers unanimously agree that the importance of operational flexibility stems from the dramatic change that has taken place since the 1980th in the market places in virtually all sectors of industry. As markets saturated, firms started to compete on product differentiation and product innovation. As a result, both the number of product variants simultaneously offered to the market, and the new product introduction rate, substantially increased. Product innovation and new product development became important research fields and concurrent engineering was developed as a new approach to structuring and managing the product innovation process, aimed at delivering better quality products at lower costs in shorter time (Clark and Wheelwright, 1993). At the beginning of the 1990th "time-to-market" was an established concept in industry, indicating the industry-wide awareness of innovation speed as a competitive weapon (Stalk and Hout, 1990). However, during the same period, firms that decided to compete on product diversity and product innovation were faced with the downside of this strategy, being a large increase in demand uncertainty at the product variant level.

For each of the product variants offered production capacity has to be reserved in order to be able to deliver the product up to a certain level, and work-in-process and inventory has to be built up. This upfront investments have to be made for setting up a supply chain for a product. These investments have to be recovered by the revenues from sales.

The initial lack of flexibility to adapt the supply chain to emerging demand for product varies to frequently led to lost sales for some product variants and product markdowns, writing-off of excess inventory for other product variants. Thus, supply flexibility became am important feature of industrial production systems. In supply chains that experienced an imbalance between demand uncertainty and supply chain flexibility, manufacturers and suppliers were put under high pressure to increase their flexibility.

Flexibility in manufacturing mainly stemmed from three sources. First, there is the variety of the manufacturing technologies employed. Technological variety allows from a large variety of different products to be made.

Second there is the flexibility in the amount of capacity available from production. Available capacity limits the volume of products that can be delivered to the market. Third there is the flexibility in the timing and frequency of the system. Timing and frequency of use are often restricted on economic grounds (due to change-over and set-up effects). Inflexibility of the timing and frequency of use of production systems lead to high levels of work-in-process and inventory, and lead to long lead times for introducing new product variants.

In the context of this we only deal with three sources of flexibility: mix flexibility, volume flexibility, flexibility in the timing and frequency of production.

We distinguish three streams of research on flexibility. First, research has been performed that investigates the flexibility concept itself. Important contributions have been delivered in this stream by Slack (1987), Barad and Sipper (1988), Sethi and Sethi (1990), Brown et al. (1984), Hyun and Ahn (1992), Gupta and Buzacott (1989), Brill and Mandelbaum (1989), Gerwin (1993), De Groote (1994), Koster and Malharte (1999), Beach et al. (2000) Pereira and Paulré (2001) and D'Souza and Williams (2000). Second, research has been performed on the evaluation of the value of specific types of output flexibility. Important contributions have been delivered in this stream by Nelson (1967, 1970, 1973), Weeks (1983, 1984), Mandelbaum and Buzacott (1990), Boyer and Leong (1996), Jordan and Graves (1995), Fine and Freund (1990), Andreon (1990), Nagarur (1992), Chandra and Tombak (1992), Gaimon and Singhal (1992), Abdel-Malket and Wolf (1991), Bernardo and Mohamed (1992), Malhotra et al. (1993), Malhotra and Ritzman (1990), Malhortra and Kehr (1994), Naudkeolyar and Christy (1992), Park and Brobowski (1989), Treleven and Elvers (1985), Van Mieghem (1998), Smunt and Meredith (2000).

Third there is the stream of empirical research on the measurement of flexibility, and on the effectiveness of flexibility in manufacturing systems. Important contributions in this stream have been delivered by Gustavson (1984), Browne et al. (1984), Swamidass and Newell (1987), Gupta and Goyal (1989), Gupta and Somers (1992), Kochikar and Narendran (1992), Dixon (1992), Fiegerbaum and Karnari (1991), Gupta (1993), Parthaserthy and Sethi (1993), Ettlie and Penner-Hahn (1994), Gupta and Somers (1996), Suarez et al. (1995, 1996), Upton (1995, 1997), and Boyer (1999).

The rest of this report is organized as follows.

Section 2 deals with conceptual research on manufacturing flexibility and serves to position the contingent and multidimensional character of the flexibility concept. Section 3 deals with results of model-based research on flexibility. First a selection of papers is discussed that deal with the problem of investing in flexible production capacity to cope with uncertainty about future demand levels for products; i.e. to create product volume flexibility and product mix flexibility. Second, a selection of papers is discussed that deal with the effects of machine, routing, and labor flexibility, and the control of these flexibilities, on the time-flexibility of production i.e. the responsiveness of production to short term demand. Finally, in Section 4 a selection of papers is discussed that report on empirical research on manufacturing flexibility. These papers provide information about the use of flexibility as a strategic option, and relationships between flexibility and observed performance. Section 5 concludes the report.

2. Conceptual research on manufacturing flexibility

A comprehensive definition of flexibility is given in Upton (1994), who characterizes flexibility as "the ability to change or react with little penalty in time, effort, cost or performance". Change may take place in different fields and may pertain to different aspects of a firm's environment or its internal processes. Moreover change may come at different levels of magnitude, and with different levels of surprise. De Groote (1994) states that "A particular technology is said to be more flexible than another, if an increase in the diversity of the environment yields a more desirable change in performance, than the change that would be obtained with the other technology under the same conditions". Thus flexibility is contingent upon the environmental diversity that one wants to cope with, and on how one prefers the one output of the system over the other. The implication is that the flexibility of a specific technology can only be evaluated in the context of a particular environment and a particular output preference.

Gerwin (1993) distinguishes four types of market uncertainty: market acceptance of kinds of products, length of product life cycle, demand for specific product options, and aggregate product demand, and two type of process uncertainty: resource availability and material availability. He relates each of these types to a different dimension of flexibility. These

flexibility dimensions are: mix flexibility, product innovation flexibility, product modification flexibility, volume flexibility, process routing flexibility and material flexibility.

As noticed by Slack (1993), each type of flexibility has two aspects: range and time. One manufacturing technology is more flexible than another if it can handle a wider range of possibilities. A manufacturing system is also more flexible than a second if it can attain a new possibility in the range in a shorter period of time (or at lower costs for the same period of time). The time and cost dimensions of flexibility are also denoted as "mobility" (Upton (1994), D'Souza and Williams (2000)).

Manufacturing flexibility requires investments and therefore it should be carefully investigated which are the uncertainties that a firm faces. At the strategic level a firm has to decide how to cope with these uncertainties; on of the possibilities being not to invest in flexibility at all. In this context Gerwin (1993) states that "an unintentional bias exists in favor of recommending more flexibility than is economically appropriate". Gerwin distinguishes four generic strategies for coping with uncertainty, being:

1) reduction of uncertainty by investing in variance reduction such as long term contracts with customers and suppliers, preventive maintenance and total quality control, and design for manufacturing

2) banking, that is, the use of flexibility to accommodate known types of uncertainty such as surges in demand

3) adaptations, that is, the defensive use of flexibility to accommodate unknown uncertainties

4) redefinition, that is, the pro active use of flexibility to raise customer expectations, to increase uncertainty for rivals and gain competitive advantages.

Each of these strategies may lead to a required flexibility at the systems level, which in turn results in a required flexibility at the operational level in combination with methods for delivering flexibility (such as organization of work and information systems), and operations control methods.

The relationships in detail between strategic choices and flexibility options are still unclear and seem to depend on type of industry (mass production, batch production, one of a kind production) and the innovativeness of the markets operated on. However, efforts have been

made to identify the qualitative relationships between technological choices and the effects for the flexibility at the various flexibility dimensions. Such taxonomies of the flexibility concept have been developed by Browne et al. (1984), Sethi and Sethi (1990), Carrea (1994), and Gupta and Goyal (1989). Koster and Malhorta (1999) provide a theoretic framework for analyzing the dimensions of manufacturing flexibility. In their paper they present the set of definitions of flexibility dimensions as shown in Table 1 and discuss the causal relationships between these flexibility dimensions. Each of the dimensions in Table 1 is characterized by a range of values that can be attained without incurring high transition penalties or large changes in performance outcomes. However as noticed by Slack (1983) and Lipton (1994), transition costs do not eliminate flexibility but inhibited the speed of change, which they consider to be the second aspect of a flexibility dimension. Thus each of the flexibility dimensions in Table 1 can be further characterized by the transition costs or speed of changing from one state to another state within the range.

Building on previous research Koster and Malhorta (1999) conclude to the following relationships between the various flexibility dimensions:

The basis of the flexibility of a manufacturing system is formed by the machine flexibility, the labor flexibility and the materials handling flexibility. Machine, labor and materials handling flexibility all pertain to the range of processing routings that can be used to make a product; a process routing being a sequence of operations on machines that lead to a product.

Table 1

Definitions of flexibility dimensions

Dimension	Definition
Machine flexibility	the number and heterogeneity (variety) of operations a machine can execute without incurring high transition penalties or large changes in performance outcomes
Labor flexibility	the number and heterogeneity (variety) of tasks/operations a worker can execute without incurring high transition penalties or large changes in performance outcomes
Material handling flexibility	the number of existing paths between processing centers and the heterogeneity (variety) of material which can be transported along those paths without incurring high transition penalties or large changes in performance outcomes
Routing flexibility	the number of products which have alternate routes and the extent of variation among the routes used without incurring high transition penalties or large changes in performance outcomes
Operation flexibility	the number of products which have alternate sequencing plans and the heterogeneity (variety) of the plans used without incurring high transition penalties or large changes in performance outcomes
Expansion flexibility	the number and heterogeneity (variety) of expansions which can be accommodated without incurring high transition penalties or large changes in performance outcomes
Volume flexibility	the extent of change and the degree of fluctuation in aggregate output level which the system can accommodate without incurring high transition penalties or large changes in performance outcomes
Mix flexibility	the number and variety (heterogeneity) of products which can be produced without incurring high transition penalties or large changes in performance outcomes
New product flexibility	the number and heterogeneity (variety) of new products which are introduced into production without incurring high transition penalties or large changes in performance outcomes
Modification flexibility	the number and heterogeneity (variety) of product modifications which are accomplished without incurring high transition penalties or large changes in performance outcomes

– Operation flexibility and routing flexibility are conditional on the machine flexibility, the labor flexibility and the materials handling flexibility, but are also influenced by product design and process design. Thus, product and process design can be important sources of manufacturing flexibility. Also, given the routing flexibility and the operations flexibility desired, machines, labor, and materials handling systems can be chosen in order to achieve this routing and operations flexibility.

– Given the types of numbers of machines, labor and materials handling systems available, and given the process routings for the products, a shop has a certain volume flexibility and a certain mix flexibility. Automation levels, labor contracts, work organization and management techniques are important factors for achieving volume flexibility and mix flexibility. These form what Gerwin (1993) calls the "system for delivering flexibility".

– Volume flexibility and mix flexibility are dimensions of output flexibility. Two other dimensions of output flexibility are the product modification flexibility and the new product flexibility. Both are dimensions of product flexibility and refer to the range of

different products that can be made. A product modification refers to a product change which leaves the functional characteristics unchanged, but leads to an improved design; a new product will have changed functional characteristics. Both flexibilities are conditional on the machine flexibility, the labor flexibility and the materials handling flexibility, but also depend on the capabilities and work organization of the product design and the process design departments.

Apart from the range, also the speed of product flexibility is important. This refers to the time and costs needed to deliver a new product, or a modified product, to the market. Speed and costs of a product introduction are determined by the time and resources needed for the product and process design phase, and also by the time and costs required to fill the production system with work-in-process and inventory of the new product, up to the level needed to support the planned supply level in the market. Thus, a production system that can economically work at lower levels of work-in-process and inventory, can have a higher ability at the product flexibility dimension.

— The last dimension of output flexibility mentioned by Koster and Malhorta (1999) is the expansion flexibility. Expansion flexibility implies adding more or new resources (machines, labor, materials handling systems) to the production system in order to be able to expand the overall output of the system. Here also, range and mobility are relevant aspects. Expansion flexibility is created at the strategic level in a company and requires the planning and control of options for acquiring capital goods, material supply, expanding human resources, expanding subcontracting, and requires the planning and control of the financing resources needed for the expansion.

Other research on the concept of flexibility can be found in: Slack (1983), Gerwin (1987), Carlsson (1989), De Meyer et al. (1989), Gupta and Goyal (1989), Mandelbaum and Brill (1989), Ramasesh and Jayakumar (1991), Chandra and Tombak (1992), Chen et al. (1992), Dixon (1992), Gaimon and Singhal (1992), Hyun and Ahn (1992), Das and Nagendra (1993), Gupta (1993), Gupta and Somers (1996, 1992), and DeTone and Tonchia (1998).

In the next section we give an overview of model based research on effects of the dimensions of technological flexibility on performance.

3. Model based research on flexibility

Model based research on flexibility generally deals with models of production systems consisting of a number of different machines, and a number of workers. Materials handling systems are sometimes mentioned but seldomly explicitely modeled. An exception to this is the vast literature of Flexible Manufacturing systems, where the production capabilities of Integrated FMSs are often constrained by the characteristics of the connecting materials handling system. In this chapter we will not pay special attention to (integrated) FMSs.

In model based research on flexibility it is generally assumed that each machine can perform a range of operations, and each worker can work at a certain range of machines. Further, an operation is defined as a specific transformation of a specific piece of material, and a process is defined as a specific sequence of operations that leads to a specific product.

The models in this research assume that the range of processes that can be carried out by the production system determine the range of products that can be made. There is often a one-to-many relationship between products and a processes. A product is uniquely defined by its technical product specifications; however, various processes may exist in a production system that lead to the same product.

Depending on the level of automation, workers are needed for carrying out the operations. This can range from operations that are carried out by workers using a machine, to operations that are carried out by machines that operate on palatized parts that are positioned and transported by a fully automated materials handling system, only requiring a worker for loading and unloading the pallets. Whatever the level of automation, workers are needed for operating the production system. Thus each machine is characterized by the required workers skills, and the shop is characterized by the skills that each worker posseses for operating a machine.

The range of technologies

The different technologies (machines, transportation systems, knowledgable workers) available in the production system determine the range of different products that can be made. The broader the range of technologies, the broader the range of products. A machine shop mainly consisting of general purpose machine tools can clearly produce a much wider range

of different products than a specialized parts manufacturing shop of a gearbox plant. It follows that the broader the technology range, the higher the product modification and product innovation flexibility that can be supported by a production system.

A large technological range however comes at a price, since it increases the costs per unit produced by the production system. When deciding about the range of technologies of a production system, a trade-off has therefore to be made. Focusing the production system on a specific range of products, requiring only a specific range of operations and routings, allows for the selection of specialized machines, operators, tools, materials handling devices, and working methods. This leads to increased learning effects, leads to increased output for a given amount of capital investment and therefore leads to a decrease in cost per time unit produced. These are known as the benefits to be obtained from focussed factories (Skinner (1984)). The decision about the technology range of a production system has to be based on a trade-off between the benefits to be obtained from low costs per unit produced, and the risks of not finding sufficient demand for the products that can be made on the system, over the economic life time of the system. Technology range therefore primarily caters for uncertainty regarding process requirements that follow from future products modification and future product innovations.

To our knowledge no research has been reported in operations management literature on the strategic question about the range of technologies that should be available in a production system. This clearly belongs to the realm of operation strategy and not to operations management. Therefore in this subsection we only discuss research that considers the range of technologies to be given. We first discuss the flexible resource investment problem, which has recently emerged as a research subject in operations management literature.

The flexible resource investment problem

We discuss the research into the optimal use of resource flexibility in order to cope with demand level uncertainty. We restrict our discussion to the research that has been performed by Andreou (1990), Fine and Freund (1990), Gupta (1993), Jordan and Graves (1995), Boyer and Keong Leong (1996), and Van Miehghem (1998). All this research studies models with multiple products and multiple resources, where each product requires only one resource, and each resource can be used to make one or more products.

Andreou (1990) presents an investment model to calculate the dollar value of flexible resource for a two-product production system with (correlated) stochastic future demand levels per product. He considers a production system consisting of a mix of dedicated and flexible resources and calculates the option value of flexible capacity as a function of the uncertainty in demand level, the costs of dedicated and flexible capacity and the revenues from sales. The analysis shows that value of flexibility can be substantial, especially under high uncertainty in demand, and that most of the benefits to be obtained from product mix flexibility are captures by having only a certain percentage of total capacity to be flexible, the optimal percentage being dependent on the variability of the demand level and the correlation in demand levels between the two products.

Fine and Freund (1990) present a model of the cost-flexibility trade-offs involved in investing in product-flexible manufacturing resources. They formulate the problem as a two-stage stochastic program. In the first stage they make its capacity decision, before the resolution of uncertainty in demand level. In the second stage, after demand levels for products are known, the firm takes its production decisions, constrained by the first stage resource decisions. They consider a situation with n different product families with for each product family a dedicated resource, and one flexible resource that can be used for any of the n product families. The problem consists of deciding about the amount of dedicated and flexible resources K_j, $j = 1$, ..., $n+1$,

- assuming per resources linear acquisition costs as a function of amount of resource installed, K_j,
- assuming linear production costs as a function of amount of product produced on dedicated or flexible resource,
- assuming that revenues are strictly concave as a function of the amount of product sold,
- assuming linear, technology independent variable production costs, and,
- assuming probabilistic information about the demand level per product family.

They proof that flexible capacity should be acquired when the expected value of its best use for each realization of demand, summed over all possible realizations of demands, exceeds its costs. They also provide results for the optimal capacity levels. For the two-product family situation, a family A and a family B, they proof that an increase in capacity costs for dedicated

resource A leads to a decrease of K_A, an increase in amount of flexible resource, K_{AB}, and a decrease in K_B. This because the flexible resource substitutes for dedicated resource B as well as for dedicated resource A. Furthermore the magnitude of decrease in K_A exceeds the magnitude of increase in K_{AB}, which in turn, exceeds the magnitude of decrease in K_B. These results illustrate the complex interactions between the problem parameters, on the one hand, and the optimal levels of investments in dedicated and flexible resources, on the other hand. For two numerical cases, Fine and Freund also investigate the sensitivity of the solutions of their model to correlation and variability in demand. Their results illustrate that the need for flexible capacity increases relative to level of risks in the presence of negatively correlated demand, and is zero, regardless of the level of risk, in the presence of positively correlated demand.

Building on the results obtained in Gupta and Buzacott (1991), Gupta (1993) considers the situation where N products with uncertain demand levels are to be made on M resources, where each of the resources has the same capacity, Q, and can process at most $1 \leq K \leq N$ different products. He develops a two-stage stochastic programming formulation of the problem to determine the optimal levels of M, Q and K, as a function of uncertainty in demand, sales revenues and resource costs.

Analysis of a number of numerical cases suggests that for this model the relative benefit of flexibility depends on scale economics. With certain cost structures, duplication of resources with limited flexibility might provide just as good an ability to cope with uncertainty in demand as investments in flexible resources.

Jordan and Graves (1995) investigate how to configure resource flexibility. They state that (1), limited flexibility, configured in the right way, yields most of the benefits of total flexibility, and (2), limited flexibility has the greatest benefits when configured to chain products and resources together to the greatest extent possible. Based on a planning model for assigning production of products to resources, they demonstrate that, for realistic assumptions about demand uncertainty, limited flexibility configurations have sales benefits that are approximately equivalent to those for total flexibility. They provide the following guidelines for identifying the best way to add flexibility:

−   try to create a circuit that encompasses as many resources and products as possible

- try to equalize the number of products to which each resource in the chain is directly connected

- try to equalize the number of resources to which each product in the chain is directly connected.

At a 10-product, 10-resources example, they illustrate the impact of complete one-chain resource flexibility on sales. For total capacity equal to expected demand, going from no flexibility to one-chain flexibility, expected sales increased by 11.8%. Even when total capacity is 25% above or below expected demand, total expected sales increases by more than 5%. This result indicates that capacity needs only roughly to be in balance with expected demand in order to get sales and utilization benefits from adding flexibility.

Jordan and Graves (1995) also develop a simple measure for the inflexibility in a given product-resource configuration. The measure, $\pi(M^*)$, is defined as the maximal probability over all groupings of products (M) that there will be unfilled demand for a group of products, while simultaneously there is excess capacity at resources making other products. The measure indicates whether adding more flexibility to the configuration is likely to lead to higher expected output.

Figure 1 taken from Jordan and Graves (1995), shows how $\pi(M^*)$ varies with demand uncertainty (μ / σ) and the number of products and resources, n, for the fully chained configuration where each resource can produce two products ($h = 2$).
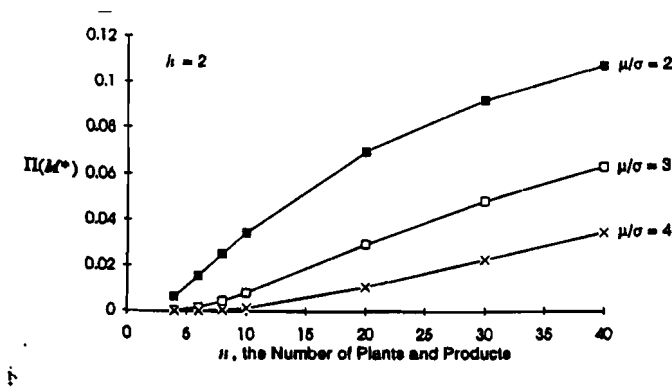

Figure 1. Flexibility Measure vs. Number of Plants and Products

The figure suggests that for high levels of demand uncertainty, and many products and resources, limited resource flexibility with only two products per resource, may not provide the same ability to cope with demand uncertainty as total flexibility.

Fig. 2, also taken from Jordan and Graves (1995), shows how $\pi(M^*)$ varies with h, the number of products that can be produced per resource, for a fully chained configuration with 40 products and 40 resources, and $(\mu/\sigma) = 2$. The figure shows that even with high demand uncertainty and many products and resources, limited flexibility (not more than 4 products per resource) can provide almost all of the benefits of total flexibility.

$n = 40$
$\mu/\sigma = 2$

$\Pi(M^*)$

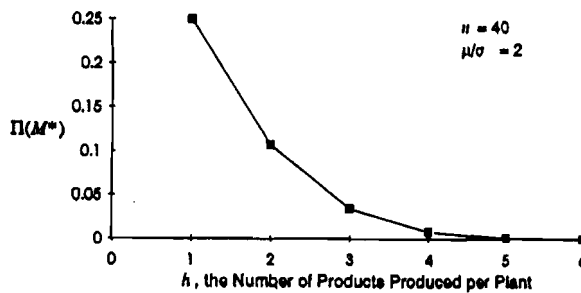h, the Number of Products Produced per Plant

Figure 2. Flexibility Measure vs. Products per Plant

Boyer and Keong Leong (1996), expand the model used by Jordan and Graves (1995), to also include the costs of changing over a resource from one process to another. Change-over costs are modeled as a loss of capacity; a percentage of available capacity is lost if the resource is used for two products. They develop a binary integer programming formulation of the problem of maximizing the expected output under stochastic demand levels for the different products, and for a given configuration of resources. For two case problems, taken from the automobile industry, the model is used to investigate the effect of different levels of resource flexibility and change-over costs on expected output. Their results indicate that for change-over costs less than 50% of available capacity, the decrease in expected output is roughly linear with in change-over costs. However, even for change-over costs up to 100%, configuration with resource flexibility have a higher expected output than configurations without resource flexibility. This is because with resource flexibility, it is still possible to decide which of the products is going to be made on a flexible resource, which is not possible

with inflexible resource. The results furthermore confirm that it is rarely beneficial to pursue total flexibility, since limited flexibility, if configured in one product-resource chain, offered approximately 95% of the output benefits of total flexibility.

Van Mieghem (1998) studied the optimal investment decisions in flexible manufacturing resources as a function of product margins, investment costs and multivariate demand uncertainty. He considers a two product firm that has the option to invest in product-dedicated resources and/or in a flexible resource. He models the problem as a two stage multi dimensional newsvendor problem. First the firm must decide on a non-negative vector of resource capacity levels, $K \in \Re_+^3$, before the product demand vector, $D \in \Re_+^3$, is observed. After demand is observed the firm decides on production quantities per resource

$x = (y_1, y_2, z_1, z_2) \in \Re_+^4$, where $y_j + z_j$ is the total amount produced of product j, y_j is

the amount of product j produced on the dedicated resource, and z_j is the amount of product

j produced on the flexible resource. The firm chooses its production vector x so as to

maximize operating profit

$$\max_{y, z \in \Re_+^2} p_1(y_1 + z_1) + p_2(y_2 + z_2)$$

subject to:
$$y_1 \leq k_1$$
$$y_2 \leq k_2$$
$$z_1 + z_2 \leq k_3$$
$$y_1 + z_1 \leq D_1$$
$$y_2 + z_2 \leq D_2$$

where $p \in \Re_+^2$ is a prize vector.

It is assumed that D is a continuous random vector that has a joint probability density function, g, which is positive over its support. The investment costs are linear in the capacity $C(K) = cK$ where $C \in \Re_+^3$ is a vector of marginal investment costs.

The investment decision is modeled as:

$$\max_{K \in \mathfrak{R}^3_+} V(K) = E\pi(K,D) - C(K)$$

where $E\pi(K,D)$ is the expected value of the operation profits.

Mathematically analyzing the properties of this problem, Van Mieghem shows how optimal investment depends on costs and prices. In particular, he derives the conditions under which it is optimal to invest only in dedicated resources, under which condition it is optimal to invest in one dedicated resource and in the flexible resource, and under which condition it is optimal to invest in all three resources. Furthermore he shows that the optimal value $V^*(K)$ is a non increasing convex function of the price vector, p. He also studies the effects of the parameters of demand uncertainty in the optimal investment. For both perfectly positively and perfectly negatively correlated demand he derives the conditions regarding marginal prices and costs under which it is optimal to only invest in dedicated resources, in one dedicated resource and the flexible resource, and in all three resources. He shows that price conditions exist under which it is optimal to invest in flexible resource, also under positively correlated demand.

The literature on the flexible resource investment problems discussed above deals with the effect of flexible resources on the demand levels that can be served over a certain period of time, with a specific configuration of dedicated and flexible resources. Demand per product and products per resource are modeled as scalar variables, neglecting the temporal manifestations of demand during the period, and neglecting the problem of following up in an on-line mode the demand per product as it unfolds over time. In the presence of change-over costs or change-over time between products on a flexible resource, the resource investment problem should be extended with the decision on the production batch size, and the total costs should also include change-over costs, inventory costs and work-in-process costs. To our knowledge such models have not been studied yet. However, much research has been performed on the effects of resource flexibility routing flexibility, and change-over time on the production order throughput time and by work-in-process. In the next section we discuss this timing aspect (or the mobility aspect) of flexibility.

Manufacturing flexibility, production order throughput time and demand responsiveness

In this subsection we discuss the research on the effects of machine, labor, and routing flexibility, on production order throughput time also in relation to the costs and capacity losses incurred by changing over a resource from processing one product to another product. In the literature in this line of research, the general assumption is made that demand per product can be modeled as a stationary stochastic variable with a given mean. Demand manifests itself over time as a random variable, either as a random interarrival time between arrivals of demand for a product, or as a random amount of products demanded per short term time period. Also it is assumed that the capacity of the production system is given, and that capacity is larger than the capacity needed to serve the average demand. Thus, in the long run, all demand can be served, and the flexibility of the production system mainly serves to make the system responsive to the short term variations in demand. It should be noted that for some production systems, demand responsiveness can also be created by keeping stocks. Thus, in such production systems, both resource flexibility and stocks are means to achieve demand responsiveness.

The effects of resource and routing flexibility on demand responsiveness has been amply studied in operations management literature. In the chapter we discuss a selection of papaers from this literature. Specifically we discuss research performed by: Wayson (1965), Nelson (1967, 1970), Fryer (1973, 1974, 1975), Treleven and Elverse (1985), Porteus (1985), Van der Veen and Jordan (1989), Malhotra and Kehr (1994), Malhotra and Ritzman (1990), Malhotra, Fry, Kehr and Donohue (1993), Jordan and Graves (1995), Kehr (2000) and Garavelli (2001). Related research results can be found in: Weeks and Fryer (1976), Günther (1981), Park and Bobrowski (1989), Treleven (1989), Corbey (1991), Park (1991), Benardo and Mohamed (1992), Bobrowski and Park (1993), Felan et al. (1993), Handkeolyar and Christy (1993), Wismer and Pearson (1993), Akmadi and Wurgaft (1994), Benjaafar (1994), Hutchinson and Pflughhoeft (1994), Fry et al. (1995), Benjaafar and Ramakrishnan (1996), Boyer and Keong Leong (1996), Daniels et al. (1996), Ho and Moody (1996), Jensen et al. (1996), Das et al.

(1997), Shafr and Charnes (1997), Kehr et al. (1999), Jensen (2000), Peters and McGinnis (2000), Smunt and Meredith (2000), Garg et al. (2001), and Nam (2001).

Most research on this subject uses systematic computer simulation as investigation tool. A short description of the simulation model used will be given for each study.

One of the first to research the effects of resource flexibility on demand responsiveness was Wayson (1965) who performed a simulation study of a simple nine-machine job shop. The shop consists of nine work centers, each containing one machine. Orders arrive according to a Poisson process. Upon arrival of an order, an order routing is generated, such that each work center has equal probability of being the first work center to be visited. After each completion of an operation in a work center, the order has to be processed on one of the other work centers with a probability of 1/9 for each of these other work centers, or is ready and leaves the shop. The model assumes zero transportation times, no labor constraints, and 100% resource availability. At all work centers, processing times are negative exponentially distributed with the same parameter. The order arrival rate is such that the shop faces a 90% utilization rate. Resource flexibility is modeled as follows. It is assumed that there is a probability that an operation can be performed at one or more other work centers. This is expressed as a real variable, z, $0 \leq z \leq 8$ where for instance $z = 2.4$ means that each operation can be performed at least 2 other work centers, and that there is a probability of 0.4 that an operation can be performed at 3 work centers. Thus $z = 0$ means no resource flexibility; $z = 8$ means total flexibility. Upon completion of an operation of an order at a work center, the next operation of the order has to be performed. The next work center is selected from among the set of work centers where the next operation can be performed. The work center is selected that, at that time, has the least number of orders in queue. At each work center, orders are processed in order of arrival time. It is assumed that production in alternative work centers is equally efficient as production on the preferred work center.
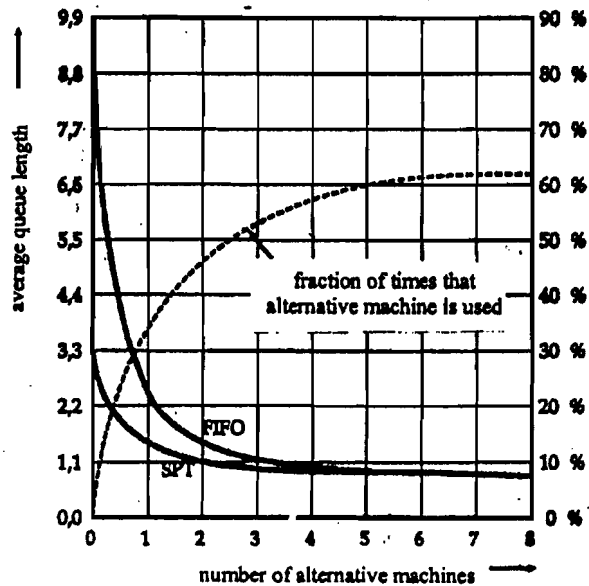
Figure 3. Average queue length as a function of alternative machines

Performance is measured with the average order throughput time. Figure 3 shows the average order throughput time as a function of the resource flexibility, z. The results show the strong impact that resource flexibility can have an average order throughput time. If each operation can be performed at two work centers, the average throughput time is 3.4 time units, as compared to 9.8 time units without resource flexibility. Even if there is only a 40% probability that an operation can be performed on an alternative resource, the average throughput time goes down from 9.8 time units to 5.4 time units. Figure 3 also shows the strong decrease in marginal benefits to be obtained from an increase in resource flexibility; there is hardly any improvement in throughput time if the flexibility is increased from 3 alternative machines to 8 alternative machines; in other words, if well-configured, most of the benefits of flexibility can already be obtained with a little flexibility. The dashed line in Figure 3 shows the fraction of times that an alternative resource is used for processing of an operation. We can see that for a flexibility of 0.4, an alternative machine is used in about 20% of the cases, resulting in the already mentioned order throughput time of 5.4 time units. The results of Wayson were obtained for a model that only captured a few elements of real life production systems. For instance, it was assumed that alternative resources are equally efficient as the preferred resource; the costs of exercising flexibility, such as forgetting and

18

relearning are neglected; it was assumed that machine capacity is the only limiting resource; and it was assumed that change-over costs and change-over times are zero.

Nelson (1967) was one of the first to study the use of labor flexibility in labor ánd machine limited production systems. Using computer simulation he investigated a two work center job shop with two identical machines per work center. The shop had characteristics in terms of arrival times, processing times and routings similar to the model used by Wayson. Nelson varied the design of the system by studying the system with 1.2, 3 and 4 workers, where each worker could work with equal efficiency in each work center. He studied centralized control, where each worker after completing his job returns to a central pool to be allocated to his next job, and decentralized control where each worker remains to work at his current work center until he runs out of work and then goes to the other work center, if work is available there. Three queue disciplines were used, First Come First Served, First in System, First Served, and Shortest Operation Time, in combination with five labor allocation rules, among which "random" and "most work in queue". The results of the simulation study indicate that labor flexibility can strongly decrease the mean and variance of the order throughput time, and that the magnitude of the effect depends on the labor allocation procedure and queue discipline used. Centralized labor allocation performs consistent better than decentralized labor allocation. Labor flexibility is only effective if the number of machine is larger than the number of workers.
Centralized worker allocation is expecially important in case workers have different efficiencies in different work centers. Workers then should preferably work in the work center were they are most efficient, unless no work is available there, and should return to this work center as soon as sufficient work is available there. Labor transfer costs (change-over time, forgetting) however, would in turn limit the frequency of worker transfer (Nelson (1970)).

Fryer (1973, 1974, 1975) investigated the effect of various labor allocation rules on order throughput time, for a three department production system, with each department consisting of four work centers with two identical machines. The production system has 12 workers who all can work with equal efficiency on all machines. Order arrival times, order routings and order processing times are all random variables with parameters such that the average worker utilization rate was 90%. Fryer distinguished inter-departmental labor (re)allocation from

intra-departmental labor (re)allocation, and also studied the effect of a (re)allocation delay on the effectiveness of labor flexibility. Workers were fully flexible over all work centers. He found that interdepartmental flexibility has a pronounced impact on performance, as opposed to intra-departmental flexibility, and that the average order throughput time could decrease with about 40%, as compared with the reference situation where to each of the 12 work centers one worker is allocated who is never reallocated. This result was obtained with a zero reallocation delay. He also found that the decrease in average order throughput time strongly depends on the reallocation delay. A delay of one time the average operation processing time still results in the average flow time decrease of 23%. However, if it takes an amount of time equal to two times the average operation processing time to reallocate a worker after his current work center has become idle, only a decrease in throughput time of 3% remains. Thus, labor (re)allocations should be fast (or pre-planned) in order to be effective.

Treleven and Elvers (1985) investigated the effect of 11 different labor reallocation rules (where to allocate) on mean and variance of queue time, mean and variance of lateness, percentage of late jobs and total number of labor transfers, for a 9 work center job shop with two machines per work center, 12 or 9 workers and random order arrivals, random routings and random operator processing times, with parameters set such that labor utilization was 90%. Workers were equally efficient in work centers were they could work. Statistical analysis of a comprehensive simulation study of this model revealed no significant differences in performance between the shop performance under any of the eleven labor allocation rules, expect for the performance criteria "total number of labor transfers". Thus the decision about where to allocate a worker seems to have hardly any impact on the order related shop performance measures. Therefore the allocation rule should be chosen that minimizes the number of reallocations, since reallocations come at a cost. The best rule in this respect seems to be to allocate a worker to the work center with the longest queue.

Park and Bobrowski (1989) investigated the effect of order release mechanisms on the effectiveness of labor flexibility in a simulation study of a 5 work center, 2 machines per work center, 5 workers job shop. They considered centralized and decentralized labor (re)allocation, for four different levels of worker flexibility. Their results indicate that the order release mechanisms has no significant impact on the effectiveness of worker flexibility

on performance, and confirmed the earlier findings regarding the strongly decreasing marginal value of increasing flexibility.

Malhotra and Ritzman (1990) investigate the environmental factors of a production system that determine the effectiveness of using resource flexibility to improve the time-to performance of the system. They use three different shop configurations, each containing a number of fabrication shops and one or two assembly shops, to represent different levels of machine flexibility and labor flexibility. They investigated the impact that yield uncertainty, capacity tightness and lot sizes have on the effectiveness of using machine flexibility to achieve a good shop performance in terms of customer satisfaction, work in process and inventory, and labor costs, in an MRP driven production system. Computer simulation is used as a research technique. The results indicate that machine flexibility is especially helpful in environments characterized by high uncertainties, tight capacities and large lot sizes. In particular, as lot size and shop utilization increase, gains achieved by having machine or worker flexibility are high for the customer service measure; with small lot sizes and low shop utilizations, improvements are moderate. Finally, simultaneous introduction of both machine and worker flexibility improves performance only marginally above that of either machine or worker flexibility alone.

Malhotra et al. (1993) investigate the impact of learning and labor attrition on the effect of worker flexibility on performance in dual constrained job shop. They study a 6 work center shop with each work center containing 4 identical machines, and with 12 workers that had different levels of flexibility. Orders arrive randomly and visit each work center just once, whereas operation processing times are exponentially distributed. Parameters are set such that a worker utilization of 85% is achieved. Orders are assigned a due date proportional to the total work content of the order. A worker is reallocated as soon as there is no more work in the queue at his current work center, and is allocated to the work center that contains the job that has been in the system for the longest time. Finally, orders are dispatched according to the earliest due date.

Malhotra et al. investigate the cost impact of acquiring worker flexibility under two levels of the learning rate, 75% and 85%, and two levels of time required to process the first order at a work center: two times the standard operation processing time and four times the standard

operation processing time. This was combined with three levels of attrition, 0%, 8% and 16%, and six degrees of required flexibility, representing the number of work centers that a worker can work in. As performance criteria were used: the mean order throughput time, the average mean tardiness, the percentage of jobs tardy and the percentage of time workers spent in learning new tasks.

Their results indicate that attrition has a significant impact on the performance measures, where this impact was primarily present in the high learning loss environment (low learning rate and high initial processing time). They also found that for high learning rate, and for low learning rate in combination with low initial processing time, mean order throughput time decreased with increasing levels of flexibility (although at a decreasing rate, as found earlier). However, for low learning rate in combination with high initial processing time, the mean flow time increased again for flexibility levels larger than three (each worker can work at three resources). This is explained by the large fraction of time spent on learning, in this learning environment. With full flexibility, time spent on learning may go up from 5% for zero attrition to 30% for 16% attrition.

In a next study, Malhotra and Kher (1994) used the same model, the same experimental conditions and the same performance measures as in Malhortra et al. (1993) to investigate the effects of worker allocation policies in the presence of differences in efficiency when performing operations in different work centers, and in the presence of finite transfer delays. Efficiencies ranging between 0.75 to 1.00 and transfer delays of 15% and 30% of the average operation processing times are considered. They investigated centralized and decentralized decisions about when to reallocate and five rules about to which work center to (re)allocate. They found that, with zero transfer delay, centralized decision making performs best with hardly any difference between the allocation rules (which confirms the results found by Treleven and Elvers (1985). However, in the presence of transfer delays the best performance is obtained with decentralized control (which limits the number of transfers) in combination with either allocation to the most efficiency work center, or the work center with the longest queue. Allocation to the most efficient work center resulted for all conditions in acceptable low mean order flow time of about 182 in a range of 178 to 225.

In a sequel to the research of Malhotra and Kehr (1994), Kher (2000) investigated the impact of flexibility on shop performance in dual constrained job shops with simultaneous learning and forgetting effects. He used the same shop model and performance criteria as in Malhotra et al. (1994). Learning was modeled with a log-linear model; forgetting was modeled by adjusting downwards the number of the next job to be processed, with a number that was logarithmically related to the number of jobs that could have been processed during the time that has passed from the last time a job was processed at this resource. The forgetting model is parameterized with a forgetting rate. Furthermore it was assumed that for each work center, an upfront amount of training is required before the worker can start his first job at that resource. Attrition rates are expressed as a ratio to this training time. In the experimental design three flexibility policies, three forgetting rates and five attrition rates were combined. The results revealed that at a high forgetting rate (85%) acquiring and using worker flexibility has a negative impact on average work order throughput time and average tardiness. For lower forgetting rates (90% and 95%) work order throughput times and tardiness did improve when employing flexibility, even for high attrition rates. These results show that the benefits of worker flexibility are situational.

In all the research discussed above, the effects of setup time and batch sizes are not considered. Orders are taken as given from outside, and for each of the operation of an order theprocessing time is given. Furthermore, the utilization of the shop is considered as an input parameter. However, it is known that batch sizes and shop utilization have a large impact on the throughput time. Therefore when designing for time flexibility (or mobility) the batch sizes and the utilization are important design parameters, and should be considered in parallel to machine flexibility and worker flexibility.
There is an abundance of research on the impact on batch sizes and capacity utilization on work order throughput time. In most of this research queuing models are used. For an overview we refer to Suri et al. (1998), Chapter 5, and Graves et al. (1993), Chapter 6. In this paper we discuss three papers that deal with the optimal setting of capacity, capacity use and batch sizes in an economic setting in the systems design phase.

Porteus (1985) studied the problem of whether or not to invest in set-up time reduction, taking into account reduced inventory related operating costs, but neglecting other advantages such

as increased time-flexibility and increased effective capacity. He considers a single product situation with a deterministic sales rate, m, setup costs, k, unit production costs, c, fractional per unit time opportunity costs of capital, b, non financial per unit time inventory holding costs, h and fractional per unit opportunity cost of capital, i. He assumes that set-up costs can be influenced at a cost

$$a(k) = a - b\ln(k) \qquad 0 < k \le k_0$$

where k_0 is the initial setup costs.

Porteus shows that for this costs model the total relevant costs is a convex-concave function over the interval $[0, k_0]$ with a unique local minimum

$$k^* = \min(k_0, \frac{2b^2 i^2}{m(ic + h)})$$

This result implies that high volume firms should invest more in setup costs reduction than low volume firms. Furthermore, for high volume products, the optimal quantity is independent of the sales rate, and total costs is a strictly concave function of the sales rate, implying the usual economics of scale. Porteus also derives results for the simultaneous setting of optimal sales rate and optimal setup costs, for a linear relationship between sales rate and price.

Van der Veen and Jordan (1989) analyze machine investment decisions, considering machine flexibility, machine capacity, production forecasts and costs related to investment, inventory, setup, material and labor. The paper focuses on the trade-offs between machine investment and utilization decisions.
Assuming machine speed as a decision variable, and assuming demand per product as a given, they develop a method to identify the optimal number of machines, M, to produce any number, N, of different products, also accounting for product allocation decisions, as well as production lot size decisions. The method considers investment costs in machines, where

investment costs depend on machine speed and setup costs. Inventory costs and labor costs in turn depend on machine speed.

The method is demonstrated with data from a sheet metal press shop, including a sensitivity analysis for the setup time. It appears that, in this example, setup times significantly affect investment decisions and total costs. The data provided by the method indicate how much a company should be willing to invest in reducing the setup times.

Building on the results obtained by Jordan and Graves (1995), Garavelli (2001) studied the operating costs incurred by three different flexibility configurations of a production system with N resources that had to produce N product families. He considered the no flexibility configuration (NF), where each resource is dedicated to one of the product families, the total-flexibility configuration, where each resource can product each product family (TF), and the limited flexibility configuration, where each product family can be produced on two resources in a closed chain between products and resources (LF). The LF configuration has been shown to provide close to complete product demand level mix flexibility, at lower costs then the TF configuration (Jordan and Graves (1995)). Garavelli studies the differences in time-flexibility (or mobility) between these three configurations. He studies the situation where orders for product families arrive at a given rate with exponential interarrival time. All product families have the same arrival rates. All orders require exponentially distributed processing time, with the same mean, independent of the resource that processes the orders. Changing over from producing one product family to another on a resource requires a deterministic setup time. Allocation of orders to resources for the configurations TF and LF works as follows. There exists a one-to-one allocation of product families to resources, indicating per family the preferred resource. Upon arrival an order is allocated to the preferred resource of the product family it belongs to, unless the order queue of that resource exceeds a given threshold, TV. Then, the order is allocated to the resource which can process that order and has the shortest queue of orders waiting. Orders are processed at the resource in order of arrival.

This process is studied for 5 levels of order arrival rates, implying 60%, 70%, 80% and 90% resource use (excluding setup times), setup times equal to ST = 0% (which is used as a bench mark) and 30% of average order processing time, two values of the threshold, TV = 3 and 10, and two values for the number of product families and number of resources $N = 5$ and 10.

Garavelli used computer simulation to determine the performance in terms of throughput and average throughput time for each of the 80 cases. The research focuses on the effects of setup times. For the ST = 0, TV = 3, N = 5 cases, the LF and TF configurations show decreases in average throughput time ranging from about 10% for the 60% net resource use case, to about 60% for the 90% net resource use case, as compared to the NF configuration. For the ST = 0, TV = 10, N = 5 cases, these numbers ranged from about 1% tot about 38%. As may be expected, resource flexibility leads to shorter order throughput times and thus improves mobility. This drastically changes for the ST = 0.3 cases. Then for TV = 3 the throughput time for the TF configuration gets extremely large for 80% resource use and is infinite (not sufficient resource available to process all orders) for 90% resource use. The LF configuration does not show this poor performance, but performance differences between the LF and NF configuration are smaller than with ST = 0; they range from 0% to 28%. Similar performance effects were obtained for the cases with N = 10. These results demonstrate that in the presence of setup times, resource flexibility should be used with caution, and provide another argument for considering investments in setup time reduction when configuring a production system.

The work of Porteus (1985), Van der Veen and Jordan (1989) and Garavelli (2001) each reveal one aspect of the intricate relationships between decisions to be made about the resource configuration of a production system (number of machines, workforce speed, efficiency flexibility, and changeover time), on the one hand, and the performance of the system in terms of throughput, throughput time and inventory for the products to be produced on the system, on the other hand.

Research in this field is still in its infancy and the knowledge currently available is clearly insufficient to allow for a systematic approach to designing production systems for specific time flexibility. Moreover, resource slack and resource flexibility also play a role in the design of the production system for providing volume flexibility and mix flexibility regarding the future demand levels. Production system design decisions are taken long before the actual demand levels become known. The short term resource slack and short term resource flexibility result from the realization of demand levels and from short term measures that can be taken in response to these demand levels, such as adjusting prices, promotional actions, and acceptance of only a part of the demand, leading to the actual sales levels that the system

has to deal with. Using short term measures to optimally adapt resource available to sales and vice versa is another important area for future research.

4. Empirical research on flexibility

In this section we discuss a selection of papers that report on empirical research on flexibility. These papers provide information about the use of flexibility as a strategic option and the relationships between the use of flexibility of various types and firm performance.

Swamidass and Newell (1987) performed an empirical study in which they collected data from 35 manufacturers in the US machinery and machine tool industry, in order to investigate the relationships between manufacturing strategy, environmental uncertainty and performance. One of their findings was that, the greater the manufacturing flexibility, the better the performance, regardless of the type of manufacturing process used.

Ettlie and Penner-Hahn (1994) conducted a survey study in US durable goods industry in order to investigate the relationships between manufacturing strategy and the various types of manufacturing flexibility found in the plants. They selected firms that had recently introduced flexible manufacturing systems or flexible assembly systems. They found that the more flexibility is emphasized in strategic focus, the more likely a plant is to have a shorter average changeover time per part family. Also plants seem to focus on delivery flexibility by concentrating on fewer part families per changeover in the production planning.

Suarez et al. (1995, 1996) studied data from 31 Printed Circuit Board plants belonging to fourteen electronic firms in the United States, Japan and Europe. They postulated five factors: product technology, production management techniques, relationships with suppliers and subcontractors, human resource management, and product development process, to have relationships with three dimensions of flexibility: mix flexibility, volume flexibility and new product flexibility. They found that newer, more automated, processes were associated with less mix flexibility and with less new product flexibility, and with more volume flexibility. Lean production management techniques did positively correlate with mix and new

production flexibility. Close relationships with suppliers and subcontractors had a positive impact on all three flexibility dimensions. Furthermore, plants with wage structures linked to plant performance had better volume flexibility, and plants that followed design for manufacturability principles (in particular, reuse of components) had greater mix flexibility and new product flexibility.

Upton (1995) investigated the product change-over flexibility in the fifty-two plants of eleven companies in the uncoated fine paper industry, based on detailed structured interviews with managers and operators, followed up with a one-day wrap-up conference per plant. He found that most of the variance in changeover flexibility (or process mobility) across plants could be explained by the work experience of the people in the plant, and the emphasis that their managers lay on changeover flexibility. The size and the computer technology of a plant were not important determinants of its mobility; it even looked like computer integration could be detrimental to the flexibility of a plant.

Gupta and Somers (1996) developed three hypotheses regarding the relationships between strategy, flexibility and performance, and tested them on the basis of survey data collected from 269 firms from precision machinery, electrical and electronics, industrial machinery, metal products and automobile and auto part firms. Survey data were collected about the opinion of the respondents regarding strategy, flexibility and their relationships in their firm. They found that the aggressiveness dimension in business strategy is significantly related to all flexibility dimensions. Aggressive organizations report that they tend to sacrifice short term profitability for going for market share, and therefore develop various forms of flexibility to be able to respond to changing market conditions. They also find that organizations pursuing a defensive strategy tend to seek very little flexibility. They furthermore found that the application of flexible manufacturing technology impacted negatively on both growth and financial performance, and that product and process flexibility had a negative relationship with growth performance. Finally, volume flexibility was found to have a positive relationship with growth performance.

Cagliano and Spina (2002) investigated the role of advanced manufacturing technologies in achieving strategically flexible production. Strategic flexibility is defined as the ability to shift

competitive and manufacturing priorities rapidly from one set of goals to another, within the same manufacturing system. Strategic flexible production is based on multi-focusedness, process integration and process ownership. They developed six hypotheses regarding the relationships between strategically flexible production strategies, and the application of advanced manufacturing technology. The hypotheses were tested in a survey of 392 firms in 20 countries in Europe, the Americas and Japan. They found that the adoption of strategically flexible production does not correlate with the more intense use of computerized equipment or software applications. However, the higher level of organizational integration required for a complete orientation to strategically flexible production often went along with greater computer based cross functional integration. Furthermore they found a positive relationship between the use of assembly robots and manufacturing quality, and between the use of MRP-II software and improvement in manufacturing lead time. However no independent effect was found of the use of automated manufacturing technology on performance improvement; technology alone seems unable to improve manufacturing performance.

The adoption of strategically flexible production was found to drive most of the improvements in manufacturing lead time, which was further reinforced by the use of cross functional computer integration. It was also found that computer integration is the main influence on significantly higher improvements in product variety.

For a comprehensive review of empirical research on manufacturing flexibility, we refer to Vokurka and O'Leary-Kelley (2000). This paper synthesizes the body of empirical research regarding content-related issues and identifies possible avenues for future research. Furthermore the papers examine several important methodological issues regarding manufacturing flexibility research, and indicates repeated methodological problems with regard to measurement validity, measurement reliability and general design, and suggest solutions.

5. Conclusion

Wrapping up the main finding in the reviewed literature we can state that

- A little bit of resource and routing flexibility, if well configured, can achieve almost all of the benefits to be obtained from resource and routing flexibility. This goes for volume, mix as well as product flexibility.
- Transfer delays, change-over times and costs, and learning and forgetting may seriously decrease the benefits to be obtained from resource or inhibit its application. This especially affects product flexibility. Thus investments in change-over time reduction should be considered, simultaneously with investments in type of amount of resources. Moreover when considering the use of worker flexibility, the effect of its use on worker efficiency and on labor costs, should be modeled in the decision problem, simultaneous with the decisions about investments in type and amount of machine resources.
- Control policies can play an important role in the impact of resource and routing flexibility, change-over costs and time, and worker efficiency on the performance. Control policies especially affect product flexibility and efficiency given the strongly decreasing marginal contribution of flexibility to performance. Control policies can effectively be used to constrain the negative impact of change-over times and costs, and differences in labor efficiency on performance, while still realizing a large part of the benefits to be obtained from resource and routing flexibility.
- Strategic focus on flexibility seems to be an important condition for creating the resource and routing flexibility required, and for delivering the systems flexibility aimed at. Strategic focus seems to impact both managerial attitude, and worker attitude and seems mobilize the knowledge acquisition and knowledge deployment processes needed to convert potential resource and routing flexibility into output flexibility via an effective "system for delivering flexibility".

References

Abdel-Malke, L. and Wolf, C. (1991), Evaluating flexibility of alternative FMS-designs – a comparative measure, International Journal of Production Economics, 23, pp. 3-10.

Andreou, S.A. (1990), A Capital Budgeting Model for product-mix flexibility, Journal of Manufacturing and Operations Management, 3, pp. 5-23.

Beach, R., Muhlemann, A.P., Price, D.H.R., Paterson, A., and Sharp, J.A. (2000), A review of manufacturing flexibility, European Journal of Operational Research, 122, pp. 41-57.

Benjaafar, S. (1994), Models for performance evaluation of flexibility in manufacturing systems, International Journal of Production Research, 32, pp. 1383-1402.

Benjaafar, S. and Ramakrishnan, R. (1996), Modeling, measurement and evaluation of sequencing flexibility in manufacturing systems, International Journal of Production Research, 34, pp. 1195-1220.

Bernardo, J.J. and Mohamed, A. (1992), The measurement use of operational flexibility in loading flexible manufacturing systems, European Journal of Operational Research, 60, pp. 144-155.

Bobrowski, P.M. and Park, P.S. (1993), An evaluation of labor assignment workers are not perfectly interchangeable, Journal of Operations Management, 11, pp. 257-268.

Boyer, K.K. and Keong Leong, G. (1996), Manufacturing flexibility at the plant level, Omega, 24, 5, pp. 495-510.

Browne, J., Dubois, D., Rathmill, K., Sethi, S.P. and Stecke, K.E. (1984), Classification of flexible manufacturing systems, The FMS Magazine, April, pp. 114-117.

Carlsson, B. (1989), Flexibility and the theory of the firm, International Journal of Industrial Organisation, 7, pp. 179-203.

Chandra, P. and Tombak, M.M. (1992), Models for the evaluation of routing and machine flexibility, European Journal of Operational Research, 60, pp. 156-165.

Chen, I.J., Calantone, R.J. and Chung, C.H. (1992), The marketing in manufacturing interface and manufacturing flexibility, Omega, 26, pp. 431-443.

Clark, K.B. and Wheelwright, S.C. (1993), Managing new product and process development; text and cases, The Free Press, New York.

Claudio Caravelli, A. (2001), Performance analysis of a batch production system with limited flexibility, International Journal of Production Economics, 69, pp. 39-48.

Corbey, M. (1991), Measurable economic consequences of investments in flexible capacity, International Journal of Production Economics, 23, pp. 47-57.

Daniels, R.L., Hooper, J.B. and Mazzola, J.B. (1996), Scheduling parallel manufacturing cells with resource flexibility, Management Science, 42, pp. 1260-1276.

Das, S. and Nagendra, P. (1993), Investigation into the impact of flexibility on manufacturing performance, International Journal of Production Research, 31, 10, pp. 2337-2354.

Das, S.K. and Nagendra, P. (1993), Investigations into the impact of flexibility on manufacturing

Das, S.K. and Nagendra, P. (1997), Selection of routes in a flexible manufacturing facility, International Journal of Production Economics, 48, pp. 237-247.

De Groote, X. (1994), The flexibility of production processes: A general framework, Management Science, 40, 7, pp. 945-993.

De Meyer, A., Nakana, J., Miller, J.G. and Ferdows, K. (1989), Flexibility: the next competitive battle, Strategic Management Journal, 10, pp. 135-144.

DeToni, A. and Tonchia, S. (1998), Manufacturing flexibility: a literature review, International Journal of Production Research, 30, pp. 1587-1617.

Dixon, J.R. (1992), Measuring manufacturing flexibility: an empirical investigation, European Journal of Operational Research, 60, pp. 131-143.

D'Souza, D.E. and Williams, F.P. (2000), Towards a taxonomy of manufacturing flexibility dimensions, Journal of Operations Management, 18, pp. 577-593.

Ettli, J.H. and Penner-Hahn, J.D. (1994), Flexibility ratios and manufacturing strategy, Management Science, 40, 11, pp. 1444-1454.

Felan, J.T., Fry, T.D. and Philipoom, P.R. (1993), Labor flexibility and staffing levels in a dual resource constrained job shop, International Journal of Production Research, 31, pp. 2487-2506.

Fine, C.H. and Freund, R.M. (1990), Optimal investment in product flexible manufacturing capacity, Management Science, 36, pp. 449-460.

Fine, C.H. and Li, L. (1988), Technology choice, product life cycle and flexible automation, Journal of Management and Operations Management, 1, 4, pp. 372-396.

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

Fryer, J.S. (1973), Operating policies in multi-echelon dual-constraint job shops, Management Science, 19, 9, pp. 1001-1012.

Fryer, J.S. (1974), Labor flexibility in multi-echelon dual-constraint job shop, Management Science, 20, 7, pp. 1073-1080.

Fryer, J.S. (1975), Effects of shop size and labor flexibility in labor and machine limited production systems, Management Science, 21, 5, pp. 507-515.

Gagliano, R. and Spina, G. (2000), Advanced manufacturing technologies and strategically flexible production, Journal of Operations Management, 18, pp. 169-190.

Gaimon, C. and Singhal, V. (1992), Flexibility and the choice of manufacturing flexibilities under short product life cycles, European Journal of Operational Research, 60, pp. 211-223.

Garg, S., Vrat, P. and Kanda, A. (2001), Equipment flexibility vs. inventory: a simulation study of manufacturing systems, International Journal of Production Economics, 70, pp. 125-143.

Gerwin, D. (1987), An agenda for research on the flexibility of manufacturing processes, International Journal of Operations and Production Management, 7, pp. 38-49.

Gerwin, D. (1993), Manufacturing flexibility in a strategic perspective, Management Science, 39, pp. 395-410.

Graves, S.C., Rinnooy Kan, A.H.G., and Zipkin, P.H. (1993), Logistics of Production and Inventory, North Holland, Amsterdam.

Gunascharan, P., Mastikaier, P., Yh-Olli, (1993), Flexible Manufacturing Systems. An investigation for research and applications, European Journal of Operational Research, 66, pp. 1-26.

Gunther, R.E. (1981), Dual-resource parallel queues with server transfer and information access delays, Decision Sciences, 12, pp. 97-111.

Gupta, D. (1993), On measurement and valuation of manufacturing flexibility, Internatonal Journal of Production Research, 31, pp. 2947-2958.

Gupta, D. and Buzacott, J.A. (1989), A framework for understanding flexibility of manufacturing systems, Journal of Manufacturing Systems, 8, 2, pp. 89-97.

Gupta, D. and Buzacott, J.A. (1991), Models for first-pass FMS investment analysis, International Journal of Flexibility Management Systems.

Gupta, Y.P. and Goyal, S. (1989), Flexibility of Manufacturing Systems: concepts and measurements, European Journal of Operational Research, 43, pp. 119-135.

Gupta, Y.P. and Goyal, S. (1992), Flexibility trade-offs in a random flexible manufacturing system: a simulation study, International Journal of Production Research, 32, pp. 527-557.

Gupta, Y. and Sommers, T. (1993), The measurement of manufacturing flexibility, European Journal of Operational Research, 60, pp. 166-182.

Gupta, Y.P. and Somers, T.M. (1996), Business strategy, manufacturing flexibility and organizational performance relationships: a path analysis approach, Production and Operations Management, 5, 3, pp. 204-233.

Gupta, Y.P. and Somers, T.M. (1992), The measurement of manufacturing flexibility, European Journal of Operational Research, 60, pp. 166-182.

Gustavsson, S. (1984), Flexibility and productivity in complex production processes, International Journal of Production Research, 22, 5, pp. 801-808.

Ho, Y.C. and Moody, C.L. (1996), Solving all formation problems in a manufacturing environment with flexible processing and routing capabilities, International Journal of Production Research, 34, pp. 2901-2923.

Hogg, G.L., Philips, D.T. and Maggard, M.J. (1979), Parallel-channel dual-resource constrained queuing systems with heterogeneous resources, AIIE Transactions, 9, pp. 352-362.

Hutchinson, G.K. and Pflughoeft, V.A. (1994), Flexible process plans: their value in flexible automation systems, International Journal of Production Research, 32, 3, pp. 707-719.

Hyun, J.H. and Ahn, B.H. (1992), A unifying framework for manufacturing flexibility, Manufacturing Review, 5, pp. 251-260.

Jensen, J.B. (2000), The impact of resource flexibility and staffing decisions on cellular and departmental shop performance, European Journal of Operational Research, 127, pp. 279-296.

Jensen, J.B., Malhotra, M.K. and Philipoom, P.R. (1996), Machine dedication and process flexibility in a group technology environment, Operations Management, 14, pp. 19-39.

Jordan, W.C. and Graves, S.C. (1995), Principles in the benefits of manufacturing process flexibility, Management Science, 41, pp. 577-594.

Kher, H.V. (2000), Examination of flexibility acquisition policies in dual resource constrained job shops with simultaneous worker learning and forgetting effects, Journal of Operational Research Society, 51, pp. 592-601.

Kher, H.V. and Malhotra, M.K. (1994), Acquiring and operationalizing worker flexibility in dual resource constrained job shops with worker transfer delays and learning losses, OMEGA, 22, pp. 521-533.

Kher, H.V., Malhotra, M.K., Philipoom, P.R. and Fry, T.D. (1999), Modeling simultaneous worker learning and forgetting in dual resource constrained system, European Journal of Operational Research, 115, pp. 158-172.

Koste, L.L. and Malhotra, M.K. (1999), A theoretical framework for analyzing the dimensions of manufacturing flexibility, Journal of Operations Management, 18, pp. 75-93.

Kumar, V. (1987), Entopic Measurement of manufacturing flexibility, International Journal of Production Research, 25, 7, pp. 957-966.

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

Malhotra, M.K. and Kher, H.V. (1994), An evaluation of worker assignment policies in dual resource-constrained job shops with heterogeneous resources and worker transfer delays, International Journal of Production Research, 32, 5, pp. 1087-1103.

Malhotra, M.K. and Ritzman, L.P. (1990), Resource flexibility issues in multistage manufacturing, Decision Sciences, 21, pp. 673-690.

Mandelbaum, M. and Brill, P.H. (1989), Examples of measurement of flexibility and adaptivity in manufacturing systems, Journal of Operations Research Society, 40, pp. 603-609.

Mandelbaum, H. and J. Buzacott (1990), Flexibility and Decision Making, European Journal of Operational Research, 44, pp. 17-27.

Morris, J.S. and Terinze, R.J. (1994), A simulation comparison of process and cellular layouts in a dual resource constrained environment, Computer and Industrial Engineering, 26, pp. 733-741.

Nagurar, N. (1992), Some performance measures of flexible manufacturing systems, International Journal of Production Research, 30, pp. 799-809.

Nam, I. (2001), Dynamic Scheduling for a flexible processing network, Operations Research, 49, 2, pp. 305-315.

Nandkeolyar, U. and Christy, D.P. (1992), An investigation of the effect of machine flexibility and number of part families on system performance, International Journal of Production Research, 30, pp. 513-526.

Park, P.S. (1991), The examination of worker cross-training in a dual resource constrained job shop, European Journal of Operational Research, 51, pp. 291-299.

Park, P.S. and Bobrowski, P.M. (1989), Job release and labor flexibility in a dual resource constrained job shop, Journal of Operations Management, 8, 3, pp. 230-249.

Parker, R. and Wirth, A. (1999), Manufacturing flexibility. Measures and relationships, European Journal of Operational Research, 118, pp. 429-440.

Parthasarthy, R. and Sethi, S.P. (1993), Relating strategy and structure to flexible automation in a test of fit and performance implications, Strategic Management Journal, 14, pp. 529-549.

Pereira, J. and Pankré, B. (2001), Flexibility in manufacturing systems: A relational and dynamic approach, European Journal of Operational Research, 130, pp. 70-82.

Peters, B.A. and McGinnis, L.F. (2000), Modeling and analysis of the product assignment problem in single stage electronic assembly systems, IIE Transactions, 32, pp. 21-31.

Porteus, E.L. (1985), Investing in reduced setups in the EOQ model, Management Science, 31, pp. 998-1010.

Ramasesh, R.V. and Jayakumar, M.D. (1991), Measurement of manufacturing flexibility: a value based approach, Journal of Operations Management, 10, pp. 446-468.

Sanchez, R. (1995), Strategic flexibility in product competition, Strategic Management Journal, 16, pp. 135-159.

Sethi, A.K. and Sethi, S.P. (1990), Flexibility in manufacturing in a survey, International Journal of Flexibility Management Systems, 2, pp. 289-328.

Shafer, S.M. and Charnes, J.M. (1997), Offsetting lower routing flexibility in cellular manufacturing due to machine dedication, International Journal of Production Research, 35, 2, pp. 551-567.

Slack, N. (1983), Flexibility as a manufacturing objective, International Journal of Operations and Production Management, 3, pp. 4-13.

Slack, N. (1987), The flexibility of manufacturing systems, International Journal of Operations and Production Management, 7, pp. 35-45.

Smunt, T.L. and Meredith, J. (2000), A comparison of direct cost savings between flexible automation and labor with learning, Production and Operations Management, 9, 2, pp. 158-170.

Suarez, F.F., Cusumano, M.A. and Fine, C.H. (1996), An empirical study of manufacturing flexibility in printed circuit board assembly, Operations Research, 44, 1, pp. 223-240.

Swamidass, P.M. and Newell, W.T. (1987), Manufacturing strategy, environmental uncertainty and performance: a path analytic model, Management Science, 33, pp. 509-524.

Tayur, S., Ganeshau, R. and Magazine, M. (1998), Quantitative models for supply chain management, Kluwer Academic Publishers, Boston.

Treleven, M. (1989), A review of dual resource constrained research, IIE Transactions, 21, pp. 279-287.

Upton, D.M. (1994), The management of manufacturing flexibility, California Management Review, winter, pp. 72-89.

Upton, D.M. (1995), Flexibility as process mobility: the management of plant capabilities for quick response manufacturing, Journal of Operations Management, 12, pp. 205-224

Upton, D.M. (1997), Process range in manufacturing: an empirical study of flexibility, Management Science, 43, 8, pp. 1079-1092.

Van Mieghem, J. (1998), Investment strategies for flexible resources, Management Science, 44, 8, pp. 1071-1078.

Van der Veen, D. and Jordan, W. (1989), Analyzing trade-offs between machine investments and utilizations, Management Science, 35, 10, pp. 1215-1226.

Vorkurka, R.J. and O'Leary-Kelly, S.W. (2000), A review of empirical research on manufacturing flexibility, Journal of Operations Management, 18, pp. 485-501.

Weeks, J.K. and Fryer, J.S. (1976), A simulation study of operating policies in a hypothetical dual constrained job shop, Management Science, 22, pp. 1362-1371.

Wisner, J.D. and Pearson, J.N. (1993), An exploratory study of the effects of operator relearning in dual resource constrained job shops, Production and Operations Management, 2, pp. 55-68.

Yelle, L.E. (1979), The learning curve: historical review and comprehensive survey, Decision Sciences, 10, pp. 302-328.