Received March 29, 2019, accepted April 11, 2019, date of publication April 16, 2019, date of current version April 26, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2911620

Backup or Reliability Improvement Strategy for a Manufacturer Facing Heterogeneous Consumers in a Dynamic Supply Chain

HENG DU[1] AND YE JIANG[1,2]
[1]School of Management and Engineering, Nanjing University, Nanjing 210093, China
[2]School of Business, Jiangsu Open University, Nanjing 210011, China

Corresponding author: Heng Du (18795898791@163.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 71501102.

ABSTRACT This paper examines a manufacturer's supply management strategies for mitigating yield risk in a complex dynamic supply chain. Two strategies can be adopted for the manufacturer: backup and reliability improvement. Consumers may select to leave (instant consumers) or wait (delaying consumers) when they confront the manufacturer's insufficient inventory. Utilizing the method of multi-agent modeling, a manufacturer and a supplier are modeled as the intelligent agents with the reinforcement learning behavior. The study shows that: 1) when the number of instant consumers is small, reliability improvement strategy should be selected; otherwise, the manufacturer should adopt a backup strategy; 2) only when mean yield is large enough, reliability improvement strategy is the optimal choice; and 3) if yield uncertainty is small, the manufacturer should choose reliability improvement strategy; otherwise, it is suitable to use a backup strategy. In addition, when the main supplier can determine its own wholesale price, it is found that: 1) when the mean yield is small, a lower wholesale price should be designed for the main supplier, to induce higher order quantity under backup strategy; and 2) the impact of yield uncertainty on the manufacturer's supply management strategy can be changed by the main supplier's adaptive pricing behavior.

INDEX TERMS Consumer behavior, multi-agent modeling, supply chain management, yield risk.

I. INTRODUCTION

Yield uncertainty, a common phenomenon across industries, is deemed as a key risk in supply chain management. Owing to the complicated production processes and unpredictable factors like weather, the yield of a manufacturer is usually smaller than the initial production quantity [1]. A typical example is the semiconductor and electronic equipment industry. Owing to the complex manufacturing process, there is a gap between the final yield and the expected output. For instance, in the Liquid Crystal Display manufacturing industry, it is common to get yield rate less than 50% [2]. Yield risk can lead to supply shortage, even will affect the performance of the downstream firms. Therefore, it is significant to find an effective way to mitigate yield risk.

In general, *backup supplier* (*B*) and *reliability improvement* (*R*) are two widespread strategies to deal with yield uncertainty. Under strategy *B,* a manufacturer has two suppliers with the same key component: a main supplier and a backup supplier. The main supplier is prone to yield uncertainty. The backup supplier is perfect reliable. The manufacturer reserves some components in advance from the backup supplier, but firstly sources from the main supplier. Only when components from the main supplier are less than the actual demand, the manufacturer will buy from the backup supplier. Depending on this flexible sourcing advantage, strategy *B* reduces the replenishment risk caused by the single unreliable supplier, especially when unexpected supply uncertainty from upstream occurs. As an example of Nokia, its main supplier, Philips Semiconductor plant, ever suffered a fire, resulting in a shortage of key components. To mitigate the impact of the upstream risk, Nokia timely sought the alternative suppliers for replenishing goods [3]. Although strategy *B* has some advantages, it is not without shortcomings. Generally, strategy *B* makes no contribution to improve suppliers' underlying performance, such as

The associate editor coordinating the review of this manuscript and approving it for publication was Nikhil Padhi.

product quality, production reliability and cost [4], [5]. Take the drawback into consideration, numerous manufacturers take actions to improve the supplier's endogenous production reliability (i.e., strategy R) instead of strategy B. Under strategy R, manufacturers directly invest resources (such as technique, finance and personnel) for the supplier, aiming to improve the production reliability (include product quality and output quantity). A classic example of strategy R is Toyota in Japan. The Toyota Supplier Support Center collaborates with upstream suppliers to develop Toyota Production System, which increases the competitiveness of the whole supply chain [6]. However, effort may spend substantial cost, even it sometimes fails to improve the supply reliability [3], [4]. The investment risk must be examined under strategy R. Therefore, a natural question for a manufacturer is how to trade off the two strategies under yield uncertainty.

Heterogeneous consumer purchase behavior plays a crucial role in making decisions for supply chain members. When a consumer confronts stockout from a manufacturer, leaving is not always the unique behavior. Some consumers will wait until the manufacturer's inventory is sufficient, and high loyalty is their obvious features, such as consumers in the automobile industry. Therefore, there are mainly two behavior modes when a consumer cannot attain a product from one manufacturer: leaving and delay [7]. For the former consumers, they will depart the store right now (give up buying), while the latter will wait for a period of time to obtain products. The complexity of consumer behavior is captured in our model. It is also one of our goals to understand how the two behaviors affect a manufacturer's strategies.

A supply chain is a complex dynamic system consisting of several firms. On one hand, the supply chain exists in an extremely uncertain environment, where almost all external elements vary all the time. On the other hand, the optimal decision in a certain period maybe not the best choice for a firm in a dynamic situation. Traditional studies on the supply chain pay more attention to seeking an optimal solution in a static business situation [8]. However, practical supply chain members are adaptive individuals, i.e., they can adjust decisions continuously according to their long-term learning experiences in a dynamic process. To the best of our knowledge, few yield uncertainty works take account of these dynamic characteristics in a supply chain. However, this paper considers the complexity and dynamics of a supply chain system. Each member is bounded rational, who is regarded as an adaptive agent. Specifically, a manufacturer and a supplier can make decisions continuously according to external changing environment. The evolutionary process of this dynamic decision is more in line with actual cases.

We develop two models to investigate the strategies for mitigating the yield risk from the main supplier. We consider a supply chain consists of two suppliers and one adaptive manufacturer in the face of many heterogeneous consumers. There are two supply management strategies for the manufacturer to mitigate supply crisis: B or R. Based on multi-agent modeling technique, we compare the profits of two strategies to address the following issues: (1) in a dynamic (multi-period) environment, which strategy is optimal for a manufacturer? (2) how do the consumer behavior and yield risk affect the manufacturer's strategy choice? (3) what is the impact of an adaptive supplier's pricing behavior on the manufacturer' decisions? Some managerial insights are obtained. For example, when the mean quality sensitivity of consumers is small, the manufacturer should use strategy B; otherwise, the optimal strategy is R; only when the mean yield is sufficiently high, the manufacturer should adopt strategy R; if yield uncertainty is small, strategy R should be adopted, or else strategy B is the better choice, but it is not always the case when the main supplier decides its own wholesale price.

This paper contributes to the literature in several aspects. First, unlike many literatures, this paper investigates how an adaptive manufacturer makes a careful tradeoff between the backup strategy and the reliability improvement strategy to choose the optimal one under yield risk. Especially, we capture the dynamic nature of supply chain system. Especially, a manufacturer and a supplier have an intelligent learning behavior to adjust decisions during a long term. Our paper focuses on the decision dynamics of supply chain members. Second, we model complex consumer behavior and further derive the market demand from a micro-level perspective of heterogeneous consumers. Two different consumer purchase behaviors are considered, and we focus on the effect of heterogeneous consumer behavior on the manufacturer's strategy. Finally, our work examines the impact of a supplier's adaptive pricing behavior on the manufacturer's supply management strategy.

II. LITERATURE REVIEW

This paper is related to the literature of yield uncertainty, backup sourcing, reliability improvement, and multi-agent modeling.

There are extensive works studying yield uncertainty. For example, Henig and Gerchak [9], Gerchak [10], and Erdem and Ozekici [11] study an inventory issue under the random yield. Anupindi and Akella [12] consider the problem of order quantity allocation between two unreliable suppliers under three different models. Tomlin [13] and Babich *et al.* [14] consider multiple strategies to reduce the yield risk, such as inventory and dual sourcing. Dong and Tomlin [15], Cai *et al.* [16], and Guo *et al.* [17] attempt to utilize financing or contract tools to deal with uncertain supply trouble. Our research differs from above studies, because we focus on comparing two supply management strategies: the backup strategy and the reliability improvement strategy, to identify the condition under which a manufacturer should choose a particular strategy.

Backup sourcing is also related to our work. Evidently, the literature can be categorized into two main streams. The first stream addresses the issue of how to cooperate with the backup suppliers to hedge against the disruption risk [18]–[22]. There are three main choices, including

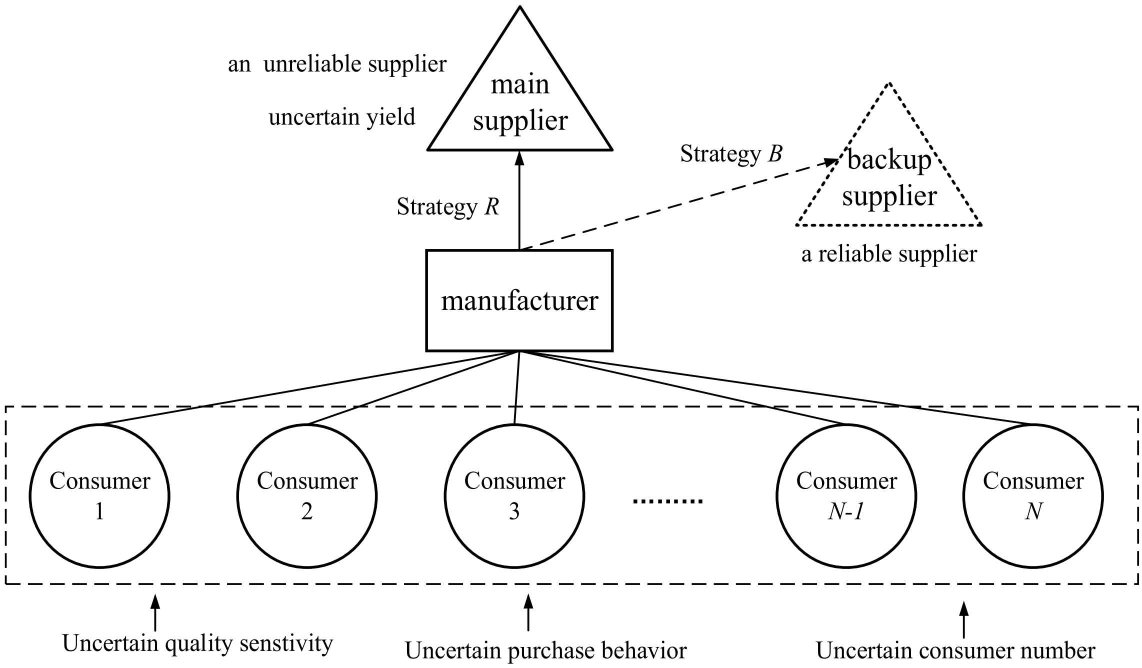


FIGURE 1. The structure of the supply chain.

advance purchase strategy, reservation strategy, and contingency purchase strategy. The other stream primarily explores various coordination contracts for backup suppliers, such as Hou and Zhao [23], Chen and Yang [2], and Chen and Xiao [24]. Nevertheless, our paper also examines the reliability improvement strategy, besides the problem of how to utilize the backup strategy.

This paper is also related to reliability improvement. The reliability improvement literature can be summarized as two types. The first type focuses on the policies for controlling the production quality to reduce the total cost [25]–[30]. Unlike the above research, we discuss how to properly adopt the reliability improvement strategy under yield risk. The second type of literature is closely related to our work, which studies the issue of a supplier's production reliability improvement. Using the methods of empirical and case analysis, Leenders and Blenkhorn [31], Krause [4] and Liu *et al.* [32] study the value of suppliers' reliability. Wang *et al.* [33] study the spillover effect on manufacturers' incentives to enhance a supplier's reliability. Wang *et al.* [34], Tang *et al.* [3], Gupta *et al.* [35], and Silbermayr and Minner [36] examine strategies (such as multiple sourcing, process improvement and a combined strategy) to reduce diverse supply risks, including random capacity, and random yield. They investigate how some external factors influence the strategy choice. Our paper is different from the second stream. The effect of reliability improvement on product quality is considered in this paper. Further, we also take account of the adaptive behavior of supply chain members. Especially, the manufacturer faces multiple uncertainties in a supply chain, who needs to make decisions based on learning experience.

Another crucial literature is multi-agent modeling (MAM). The traditional approaches of operations research and optimization are widely used in the supply chain management, such as game theory and dynamic programming. Nevertheless, practical parties among the supply chain are bounded rational [37], who could not acquire full information in a complex environment and is hard to find an optimal decision because of own ability. Actually, they often make decisions by learning the environment through past experience in most cases. Consequently, differing from the prior research, MAM is introduced to depict the supply chain agents' learning behavior. Owing to the prominent strengths (distribution, rapidness, autonomy, etc.) to cope with complex problems, MAM has become an effective and popular paradigm penetrating into the field of supply chain. With respect to MAM, many problems on supply chain have been studied, such as platform supply chain networks [38], [39], production scheduling [40], [41], and products management [42], [43].

III. THE BASIC MODEL

A. PROBLEM DESCRIPTION

Consider a supply chain consisting of two suppliers and one manufacturer. The manufacturer can order a key component from two suppliers, a main supplier and a backup supplier. The main supplier is subject to yield risk, but the backup supplier is reliable. The components provided by two suppliers are homogenous. The manufacturer sells a finished product in a market consisting of N consumers. Because of upstream supply uncertainty, the manufacturer faces two supply management strategies to mitigate the risk: the backup supplier (B) or reliability improvement (R) [34]. Fig. 1 shows the detailed economic structure.

The time sequence of the event of this model is as follows.

Stage 1: the manufacturer firstly chooses the supply management strategy: B or R.

Stage 2: under the given strategy, the second stage is made of multiple periods. During each period, the sequences of events proceed as follows:

(1) At the beginning of the period, two suppliers offer the wholesale price contract.
(2) If the manufacturer selects strategy B, he decides the *order quantity* from the main supplier and the *reservation capacity* from the backup supplier; if the manufacturer selects strategy R, he decides the *order quantity* from the main supplier and *investment effort level*. He only orders once in each period.
(3) The main supplier faces uncertain yield, and fulfills the order for the manufacturer.
(4) The manufacturer determines the *selling price*.
(5) Consumers arrive in turn and buy products from the manufacturer. The demand is fulfilled in a first-arrived-first-fulfilled principle. Especially under strategy B, if the manufacturer's order from the main supplier is less than the realized demand, he will purchase from the backup supplier to meet the insufficient demand.
(6) The manufacturer deals with the leftovers, which will become the inventory to sell in future.

Table 1 defines the key parameters and variables used throughout the paper.

In the following, we discuss the detailed behavior of each agent.

B. CONSUMER'S BEHAVIOR

Similar with [43], to capture the uncertainty of the market demand, we assume that N is a random variable following normal distribution in each period; i.e., $N \sim N(\mu_0, \sigma_0^2)$. Each consumer makes purchase decisions independently.

Two kinds of consumers are considered. When consumers cannot obtain a product, some consumers will give up buying right now, and generate a penalty cost c_p (shortage cost) for unfilled demand, i.e., instant consumers (denoted as consumers Y) here; others will select to wait until the manufacturer has sufficient inventory, i.e., delaying consumers (denoted as consumers D) [7]. In other words, unmet demand for consumers Y (D) is lost (backlogged); unmet demand of consumers D will be fulfilled firstly in next periods if the stock is sufficient. We assume that the percentage of consumers Y is α, and that of consumers D is $1-\alpha$, $0 < \alpha < 1$.

At the beginning of each period, every consumer knows the retail price. Then, each consumer chooses one product at most. Adopting the discrete choice model [5] in classical microeconomics to describe the purchase behavior, we assume that the utility of consumers is

$$U_i = \theta_i(m_0 + \Delta m) - p + \varepsilon_i (1 \leq i \leq N) \tag{1}$$

where m_0 and Δm are the basic quality and improved quality of the manufacturer's products, respectively. p denotes the retail price of the finished product. θ_i represents the sensitivity of a consumer to a product's quality. ε_i is a random variable, which indicates the impacts of those uncertain factors, such as personality, status, and career. To define the heterogeneous consumers, we assume that θ_i is a random variable following normal distributions: $\theta_i \sim N(\mu_1, \sigma_1^2)$. Similar to the literature [5], [44], we assume that ε_i follows Gumbel distribution:

$$P(\varepsilon_i \leq x) = \exp(-\exp(-(x/\varphi) - \tau')) \tag{2}$$

with mean zero and variance $\varphi^2\pi^2/6$ (τ' is the Euler's constant).

According to above objective function, consumers confirm whether to buy the product. If U_i is positive, the consumer will choose to buy one product; otherwise, the consumer will not buy any products. The detailed consumer purchase behavior is showed in Fig. 2. Therefore, unlike the traditional literature, this paper computes the total market demand d_t in each period

TABLE 1. Notations in this paper.

Superscript	Define
B	the backup strategy
R	the reliability improvement strategy
Subscript	**Define**
t	period
i	consumer index
Decision variables	**Define**
p	(selling) price of the manufacturer
q	order quantity of the manufacturer
e	investment effort level
k	reservation capacity
Parameters	**Define**
U	utility of consumers
N	number of consumers
ε	uncertainty of consumer's valuation
α	the fraction of instant consumers
θ	quality sensitivity of consumers
d	total market demand from consumers
w	the main supplier's wholesale price
π	profit
c_0	unit production cost of the raw material
c_1	unit price of reservation capacity
w_b	emergency sourcing price from the backup supplier
h	unit holding cost
β	effort cost factor
x	successful probability of reliability improvement
υ	improved quality for marginal effort
m_0	basic product quality level
Δm	improved product quality level under strategy R
c_p	penalty cost
λ	the factor on yield uncertainty
η	the committed cost
μ_0	expectation of the consumer's number
σ_0	standard deviation of the consumer's number
μ_1	expectation of the quality sensitivity
σ_1	standard deviation of the quality sensitivity
μ_2	expectation of the price sensitivity
σ_2	standard deviation of the price sensitivity
μ_3	expectation of the yield uncertainty
σ_3	standard deviation of the yield uncertainty

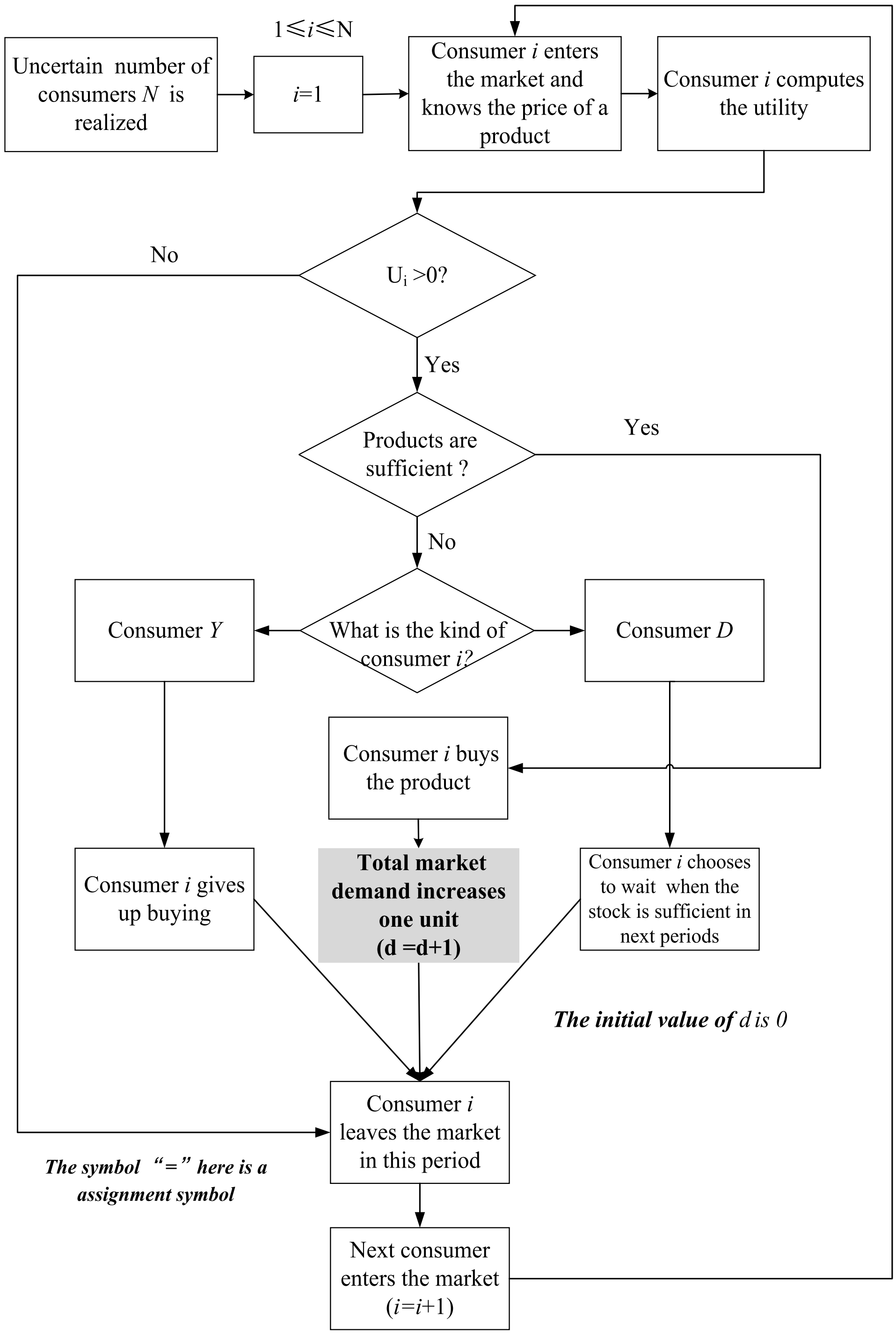


FIGURE 2. **The consumer purchase behavior and the total market demand in each period.**

by aggregating the buying quantity of all consumers in the multi-agent program following Fig. 2.

C. MANUFACTURER'S BEHAVIOR

The manufacturer can adopt two reactive approaches to reduce supply risk: strategy B and strategy R. Once a strategy has been determined, the choice cannot be changed. Because it is costly for a firm to break the agreement with partners [4].

Under strategy B, the manufacturer utilizes two suppliers to mitigate yield risk. Before each selling periods, the manufacturer orders q^B units of key components from the main supplier with a wholesale price w. One finished product is made of one unit of key component. k units of capacities are also reserved from the backup supplier. The unit price of the reservation capacity is c_1. Owing to yield uncertainty of the main supplier, the manufacturer only receives λq^B units. The common proportion model is applied to depict this random yield phenomenon. $\lambda(0 \leq \lambda < 1)$, a multiplication factor reflecting the uncertain yield, is set to be a random variable following normal distribution, $\lambda \sim N(\mu_3, \sigma_3^2)$ [34]. When a selling period starts, if λq^B is less than the actual demand $d(\lambda q^B < d)$, the manufacturer buys from the backup supplier via an expensive price w_b to meet extra demand, $w_b > w$ [21], [23]. Namely, the manufacturer sources from the backup supplier only when the demand is insufficient. Here the backup supplier is assumed to be exogenous and two suppliers' capacities are infinite.

Profit maximization is regarded as the goal for the manufacturer, who is a rational decision-maker. The objective profit function for strategy B in each period is

$$\begin{aligned} \max_{p_t^B, q_t^B, k_t} \pi_t = {} & p_t^B \cdot \min\{d_t, \lambda_t q_t^B\} - [\eta \cdot q_t^B + (1-\eta) \cdot \lambda_t q_t^B] \\ & \times w - c_0 \cdot \lambda_t \cdot q_t^B - h \cdot (\lambda_t q_t^B - d_t)^+ \\ & + I \cdot (p_t^B - w_b - c_0) \cdot \min\{d_t - \lambda_t q_t^B, k_t\} \\ & - c_p \cdot (d_t - \lambda_t q_t^B - k_t)^+ - c_1 k_t \end{aligned} \tag{3}$$

where π_t is the manufacturer's profit in each period; p_t^B, q_t^B and k_t (decision variables) are selling price, order quantity and reservation capacity, respectively. The first term is the revenue from the main supplier. The second term is the total purchase cost from the main supplier. $\eta(0 \leq \eta \leq 1)$ is the committed cost, reflecting the fact that firms sometimes incur a fraction of the procurement cost for undelivered components [45]. The third term is the total manufacturing cost. The fourth term is the total inventory cost and I is a sign index; $I = 1$ $(I = 0)$ represents that the manufacturer does (not) source from the backup supplier; $(\lambda_t q_t^B - d_t)^+ = \max\{\lambda_t q_t^B - d_t, 0\}$. The fifth term is the profit about emergence purchase from the backup supplier. The sixth term is the penalty cost of shortage. The seventh term is the total cost on reservation capacity.

Under strategy R, the manufacturer only sources from the main supplier. The main supplier is asked for collaborating with the manufacturer—that is, the manufacturer exerts investment effort (such as finance, technique and personnel) to increase supplier reliability in the production process. It is assumed that the manufacturer cannot observe the improvement result before the resource investment [34]. Additional, the improvement may fail. Here we assume the successful probability is x. When the action is taken successfully, the supplier becomes more reliable in two aspects. On one hand, the mean yield delivered to the manufacturer increases. Specifically, if the manufacturer exerts an investment effort e, the random yield factor μ_3 increases $\Delta\mu$. $\Delta\mu = \ln(1+e)$, $\lambda^* \sim N(\mu_3 + \Delta\mu, \sigma_3^2)$ [34]. On the other hand, the quality of components is enhanced. The improved quality is Δm, $\Delta m = \upsilon \cdot e(\upsilon > 0)$. υ is improved quality for marginal effort. If the action is failed, the quantity and quality both remain at original level, $\Delta\mu = \Delta m = 0$. It is assumed that effort e incurs a cost $\beta e^2/2$ for the manufacturer to improve the reliability of the main supplier [3], β is *effort cost factor*.

The objective for strategy R in each period is

$$\begin{aligned} \max_{p_t^R, q_t^R, e_t} \mathrm{E}(\pi_t) = {} & x \cdot [p_t^R \cdot \min\{d_t, \lambda_t^* q_t^R\} - (\eta \cdot q_t^R + (1-\eta) \\ & \cdot \lambda_t^* q_t^R) w - c_0 \cdot \lambda_t^* \cdot q_t^R - h \cdot (\lambda_t^* q_t^R - d_t)^+ \\ & - c_p \cdot (d_t - \lambda_t^* q_t^R)^+] + (1-x) \\ & \cdot [p_t^R \cdot \min\{d_t, \lambda_t q_t^R\} - (\eta \cdot q_t^R + (1-\eta) \\ & \cdot \lambda_t q_t^R) w - c_0 \cdot \lambda_t \cdot q_t^R - h \cdot (\lambda_t q_t^R - d_t)^+ \\ & - c_p \cdot (d_t - \lambda_t q_t^R)^+] - \beta e^2/2 \end{aligned} \tag{4}$$

where x is the successful probability of reliability improvement; $p_t^R \cdot \min\{d_t, \lambda^* q_t^R\}$ is the total revenue; $(\eta \cdot q_t^R + (1-\eta) \cdot \lambda_t^* q_t^R) w$ is the total purchase cost; $c_0 \cdot \lambda_t^* \cdot q_t^R$ is the total manufacturing cost; $h \cdot (\lambda^* q_t^R - d_t)^+$ is the total inventory holding cost; $c_p \cdot (d_t - \lambda^* q_t^R)$ is the total penalty cost; $\beta e^2/2$ is the total investment cost; $\lambda^*(\lambda)$ is the factor about yield level when the investment effort is effective (failed).

Usually, a firm is bounded rational and cannot acquire precise information about complex environment. Multiple uncertainties (yield uncertainty, demand uncertainty and consumer behavior uncertainty) exist in the supply chain simultaneously, which will increase the environmental complexity. Additional, some variables (such as price, order and inventory) about the two objective functions are dynamic, which will change according to the environment. To sum up, as a result of the complexity and dynamics, it is hard to derive the optimal decision based on traditional operations research methods. In an actual situation, the manufacturer often attempts diverse policies through learning from past experiences to obtain a better strategy. The dynamic behavior of learning from historical experiences can be well described by reinforcement learning algorithm (RL). This paper uses the method to simulate this learning process.

The reinforcement learning algorithm is firstly established by Sutton and Barto [46], which is used to search a better solution through trial and error when an agent is unknown about the external environment. One of the well-known algorithms about RL is Q-learning [47], which is adopted in this paper. This model is shown in Fig. 3.

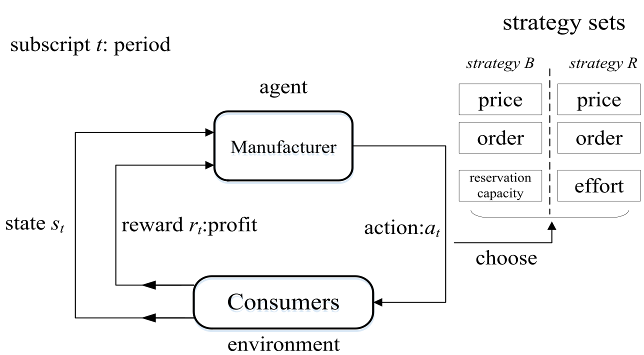


FIGURE 3. The principle of reinforcement learning model.

Four strategy sets are considered because of four decision variables: price, order quantity, reservation capacity, and effort level. Each strategy set contains finite actions. After selecting an action a (price, order quantity, reservation capacity or effort level) from a set in each period, the manufacturer gets a reward value r from environment (here denotes the profit) and changes from one state to another. Q value, based on the reward and state in the Q-learning model, is the important variable to decide the next action in subsequent periods. It is updated in each period as follows. Profit maximization is the aim for all intelligent agents.

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t)+\alpha(r_{t+1}+\gamma \max_a Q(s_{t+1}, a)-Q(s_t, a_t)) \tag{5}$$

where $Q(s_t, a_t)$ is the strength value with action a_t and state s_t, reflecting the agent's beliefs about actions, $s_t = \{s_1, s_2\}$. s_1: the profit gap is positive, i.e., $\Delta\pi = \pi_t - \pi_{t-1} > 0$; s_2: the profit gap is equal or negative, i.e., $\Delta\pi = \pi_t - \pi_{t-1} \leq 0$. r (profit) is the reward of the current action; α is the learning rate and γ is the discount factor. $\max_a Q(s_{t+1}, a)$ is the maximum value function among all actions.

To trade off exploration and exploitation behavior in RL, the traditional Softmax method is utilized to choose actions. Each action is searched with Bolzmann probability distribution.

$$\mathrm{e}^{Q(a)/\tau} / \sum_{b=1}^{n} \mathrm{e}^{Q(b)/\tau} \tag{6}$$

where $Q(a)$ is the strength of action a. τ is an internal coefficient of the algorithm, which controls the frequency of exploration and exploitation behavior.

The manufacturer agent determines each decision variable from a finite set. Specifically, $q \in [0, Q]$, $k \in [0, K]$, $e \in [0, E]$ and $p \in [P_{\min}, P_{\max}]$. $P_{\min}$ is not less than $w + c_0$; otherwise, the unit profit is negative.

Based on this learning mechanism, we have Algorithm 1 as follows.

Algorithm 1

Step 1: $t \leftarrow 1$.

Step 2: the exogenous parameters are initialized when $t = 1$, such as the holding cost, the main supplier's wholesale price and the internal variables of the RL algorithm.

Step 3: according to the RL method, decision variables are selected from a finite action set under each given strategy (B or R). Specifically, if under strategy B, the manufacturer selects the order quantity q_t^B and reservation capacity k_t according to formula (6); if under strategy R, the manufacturer chooses the order quantity q_t^R and effort level e_t.

Step 4: the manufacturer determines the selling price p_t^i according to formula (6), $i = B, R$.

Step 5: consumers in turn adopt the behavioral rule in Fig. 2 to purchase the product.

Step 6: the manufacturer deals with the inventory and computes the profit in current period t, then strength value with actions (Q value) are updated by formula (5).

Step 7: enter next period ($t \leftarrow t + 1$) and go to step 3 until termination.

Step 8: compare the average profit of the manufacturer under two strategies B and R, and then the manufacturer selects the better one.

IV. SIMULATION EXPERIMENTS AND ANALYSIS

In this section, the simulation experiments are firstly designed. Then the method of sensitivity analysis is used to investigate the effects of consumer behavior and yield risk on the manufacturer's reactive strategies.

A. EXPERIMENT DESIGN

Data in real-world from industries is preferred to utilize in simulation. However, it is not easy to obtain these secret resource in most cases, especially operations information about firms. Multi-agent modeling technique adopts massive data to perform many experiments, which avoids these limitations.

Parameters of this experiments are set as Table 2. Some parameters are set several values to reduce the effect of random values on the results of the experiments.

The manufacturer's price is limited to the range $[w + c_0, 40]$($P_{\min} = w + c_0, P_{\max} = 40$). The step size of price is 2.

TABLE 2. The values of parameters in experiments.

Parameters	Value
α	0.2,0.3,0.4,0.5,0.6
w	3,4,5,6,7
c_0	1,2,3,4,5
c_1	2,4,6,8,10
w_b	12,14,16,18,20
h	6,8,10,12,14
β	0.6,0.8,1,1.2,1.4
x	0.4,0.5,0.6,0.7,0.8
υ	0.4,0.5,0.6,0.7,0.8
m_0	10,20,30,40,50
c_p	10,12,14,16,18
η	0.1,0.2,0.3,0.4,0.5
μ_0	110,120,130,140,150
σ_0	10,15,20,25,30
μ_1	0.6,0.8,1,1.2,1.4
σ_1	0.2,0.3,0.4,0.5,0.6
μ_2	1.5,1.8,2.1,2.4,2.7
σ_2	0.2,0.3,0.4,0.5,0.6
μ_3	0.3,0.4,0.5,0.6,0.7
σ_3	0.1,0.12,0.14,0.16,0.18

The manufacturer's order quantity is limited to the range $[0, 130]$($Q = 130$). When the order quantity is 0, the manufacturer will not purchase the product. The upper bound of ordering quantity is designed large enough, to ensure sufficient replenishments during each period. The step size of order quantity is 10.

The manufacturer's reservation capacity is restricted to the range $[0, 50] (K = 50)$. The step size is 5. And the effort level is within the interval $[0, 10] (E = 10)$. The step size is 0.5.

The parameters of the RL model are designed as follows. The learning rate $\alpha = 0.5$, discount factor $\gamma = 0.8$, initial $\tau = 20$, and the decreasing rate of τ is 0.9.

Supply management strategy is a long-term decision for the manufacturer, however, market environment is more uncertain. After the supply management strategy was decided, market environment may change, i.e., the values of parameters change. A real situation is corresponding to a combination of the values of parameters. Simulation experiments are conducted on the Eclipse platform with Java programming codes. Experiments are all carried out considering all parameters with multiple values. To better illustrate the supply management strategy, we depict figures average over combinations/experiments. This method is not uncommon in the economics literature [48]. Each simulation is run 100 times with different random seeds, and each time lasts for 10000 periods to give a manufacturer abundant time to attempt different strategies. By employing the combination/experiment method, we can obtain some important results that are robust to environment change.

B. STRATEGY ANALYSIS

Fig. 4 shows the effect of learning behavior on the manufacturer's profit under strategy B. From Fig. 4, we find that the manufacturer's profit converges asymptotically to a steady state through a long time learning process. Similarly, there are the same phenomena for other variables (such as price and effort level). In the following research, the average of the manufacturer's steady profits is calculated and compared to analyze the experiment results (all the following figures in this paper are the *average* of manufacturer's steady data).

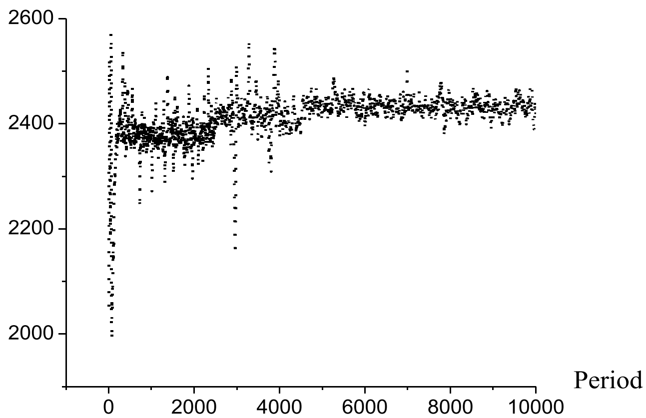


FIGURE 4. The profit evolution of the manufacturer under strategy *B*.

1) THE IMPACT OF THE CONSUMER BEHAVIOR

In this subsection, the impact of the consumer behavior on reactive strategies is explored, including the number of consumers Y and sensitivity to the quality.

First, we study the effect of the number of consumers Y on the two strategies. The value of α is set to the range $[0.1, 0.8]$. Other parameters are the same as those in subsection IV.A.

Fig. 5 illustrates the reactive strategies of the manufacturer. It shows that if the percentage of consumers Y (α) is small ($\alpha \leq A$), the reliability improvement strategy (strategy R) is better; if α is large ($\alpha > A$), the backup strategy (strategy B) is the better choice. When the number of instant consumers (α) is small, the market is made up of many delaying consumers (consumer D), who are usually loyal to a firm. Naturally, shortage cost of insufficient stock for consumers Y is not large. The manufacturer could invest more effort under strategy R, to enhance the product's quality and yield stability. A higher selling price is decided due to the higher product quality, so that the manufacturer earns higher marginal profit from these loyal consumers. Thus, it is more rewarding for a manufacturer to select strategy R if α is small. However, strategy B should be adopted when α is large. Because shortage quantity and penalty cost are larger when the number of consumers Y increases (Fig. 6). Under this situation, adequate stock for all consumers should be firstly focused on. It is more reliable for strategy B to reserve some components in advance, which could reduce yield risk.

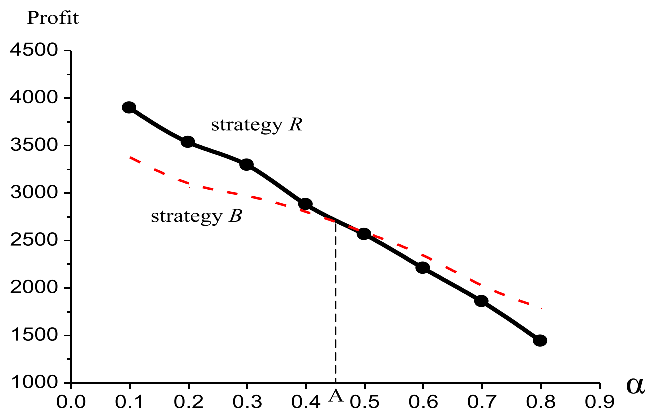


FIGURE 5. The manufacturer's profits versus the number of instant consumers.

Next, we explore how the consumers' sensitivities to the quality affect the manufacturer's strategies.

The profits of two strategies and consumers' sensitivities to the quality μ_1 are presented in Fig. 7. Fig. 7 implies that when μ_1 is less than a threshold ($\mu_1 \leq A$), strategy B is better than R; but when μ_1 is larger than this critical value, the result is opposite. What contributes to this phenomenon? On one hand, consumers' response to the product quality change is little if μ_1 is small. Products with high quality is not preferred among most consumers. Consumers are not willing to spend more money to buy a product. Therefore, effort investment under strategy R may only play a minor role to increase revenue. The profit is often offset by input cost. On the other hand, consumers' utility is enhanced much when μ_1 is large. Products about high quality are more keened on. Thus, improving product quality is a main incentive for the

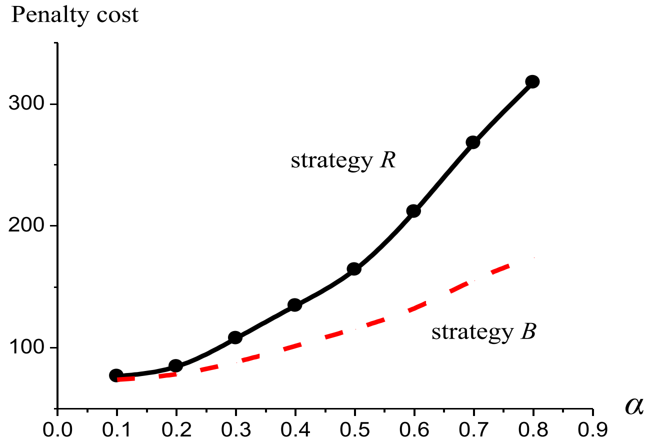


FIGURE 6. The total shortage cost versus the number of instant consumers.

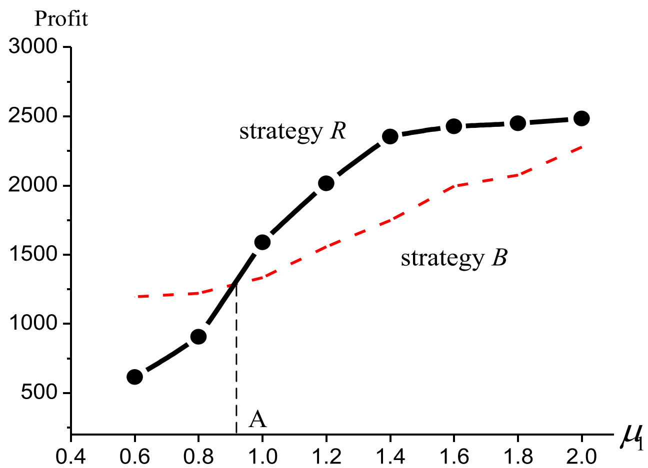


FIGURE 7. The manufacturer's profits versus consumer's quality sensitivity.

manufacturer to adopt strategy R. In brief, it is suggested to adopt strategy B when μ_1 is small. On the contrary, the manufacturer should use strategy R when μ_1 is large.

2) THE IMPACT OF YIELD RISK

The effect of the yield (expectation μ_3 and standard deviation σ_3) on strategy is investigated in this part, see Figs. 8 and 9.

As displayed in Fig. 8, the profit under strategy B is higher than that under strategy R when μ_3 is not large. But if μ_3 is large enough ($\mu_3 > A$), strategy R is superior to strategy B. As a matter of fact, penalty cost becomes a key tradeoff factor between two strategies. Owing to lower order fulfillment rate (μ_3 is not large), the manufacturer invests much effort to improve yield reliability under strategy R. However, failure possibility may affect result. It is not easy to invest resources to get ideal profit. In contrast with strategy R, market demand is partly guaranteed by virtue of reservation capacity under strategy B. Consequently, strategy B is more suitable, which could decrease penalty cost. Conversely, when μ_3 is large, most orders can be fulfilled and shortage cost is not large. Manufacturer is motivated to focus on improving product quality to get more profits. Therefore, strategy R is the optimal strategy in this circumstance. In short, strategy B should be selected when μ_3 is not larger than a threshold; otherwise, strategy R will be preferred.

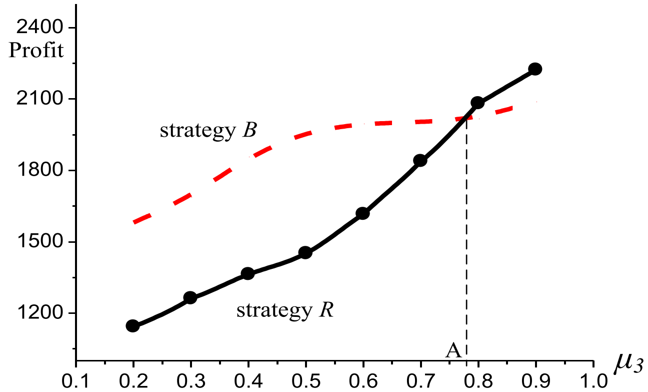


FIGURE 8. The yield expectation versus strategies.

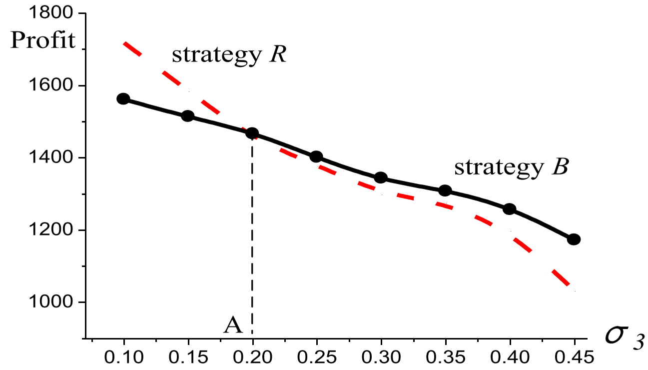


FIGURE 9. The yield uncertainty versus strategies.

From Fig. 9, we know that strategy R should be selected when σ_3 is small ($\sigma_3 < A$); otherwise ($\sigma_3 > A$), strategy B is the better choice. If the yield uncertainty is small ($\sigma_3 < A$), supply risk can be mitigated for the manufacturer. More accurate information on the main supplier will be acquired by the manufacturer, which will enhance the order fulfillment rate. Some unnecessary shortage cost and inventory cost are also saved. Additional, the main supplier's wholesale price is lower than that of the backup supplier. Hence, it is not economical to seek the backup supplier when yield uncertainty level is small. But if the uncertainty is large ($\sigma_3 > A$), it is difficult for the manufacturer to forecast order quantity received. Inaccurate information will increase supply risk, which further raises shortage/inventory cost. Thus, some components should be ordered from the backup suppler in advance, to guarantee the supply reliability. After all, supply reliability should be deemed as the most crucial factor when yield uncertainty is large. Hence, it is not fit to use strategy R, under which reliability improvement sometimes may be ineffective.

V. THE EXTENDED MODEL

The main supplier's wholesale price is assumed to be exogenous in the above research. Now we address how the manufacturer selects supply management strategy when the main supplier can decide the unit wholesale price.

A. THE MAIN SUPPLIER'S LEARNING BEHAVIOR

The main supplier is an adaptive agent who can dynamically change wholesale price, w_t, according to the past experience and external environment. Profit maximization is also the goal for the main supplier. The profit function of the main supplier is

$$\pi_t = w_t \cdot \lambda q_t + s \cdot (1-\lambda) q_t - c_2 q_t \tag{7}$$

where c_2 is the unit production cost of the main supplier, $s < c_2 < w_t$. s is the unit salvage value for the remaining materials or products, which could be remanufactured.

The algorithm process of the event is updated considering the main supplier's adaptive behavior: the whole process of Algorithm 2 is similar to that of Algorithm 1. Step 3 includes the main supplier's pricing behavior.

Algorithm 2

Step 1: $t \leftarrow 1$.

Step 2: the exogenous parameters are initialized when $t = 1$.

Step 3: if $t \leq n$, the wholesale price w_t is initialized; if $t > n$, a simple adjustment rule is utilized by the main supplier to decide wholesale price

if $\pi_{t-1} < \frac{1}{n} \sum_{j=t-(n+1)}^{t-2} \pi_j$, $w_t = w_{t-1} - \Delta w$; else if $\pi_{t-1} > \frac{1}{n} \sum_{j=t-(n+1)}^{t-2} \pi_j$, $w_t = w_t + \Delta w$; else no change is made. w_t is determined through learning past experience.

Step 4: if under strategy B, the manufacturer determines the order quantity q_t^B and reservation capacity k_t; otherwise, the manufacturer decides the order quantity q_t^R and effort level e_t.

Step 5: the main supplier computes the profit, which is regarded as the decision reference for next periods.

Step 6: the manufacturer determines the selling price p_t^i, $i = B, R$.

Step 7: consumers in turn purchase the product according to the behavior in Fig. 2.

Step 8: the manufacturer computes the profit in period t, then strength value with actions (Q value) are updated.

Step 9: enter next period ($t \leftarrow t+1$) and go to step 3 until termination.

Step 10: compare the average profit of the manufacturer under two strategies B and R, and then the manufacturer selects the better one.

Here the wholesale price adjustment rule is similar to Jiang and Sheng [49].

B. THE IMPACT OF AN ADAPTIVE SUPPLIER

New parameters of this experiment are designed in the following. $\Delta w = 0.1, 0.2, 0.3, 0.4, 0.5$; $n = 8, 10, 12, 14, 16$; $c_2 = 2, 3, 4, 5, 6$; $s = 0.5, 1, 1.5, 2, 2.5$. The other parameters are the same as those in subsection IV.A. We carry out this experiment again and analyze the result.

Several new and significant findings are presented as follows. Firstly, Figs. 10 and 11 illustrate how the yield's expectation (μ_3) affects the wholesale price, and the manufacturer's strategies. The result is inconsistent with that in the basic model (section IV.B.2)).

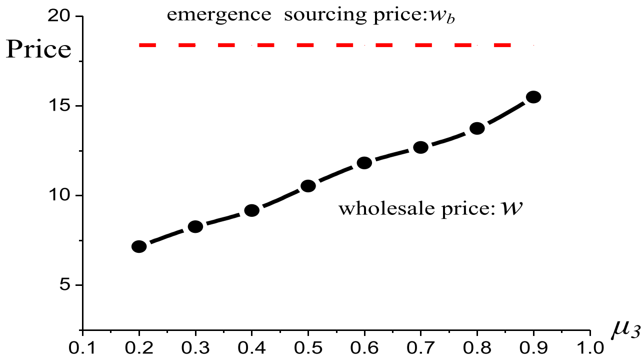


FIGURE 10. The main supplier's wholesale price versus yield expectation under strategy *B*.

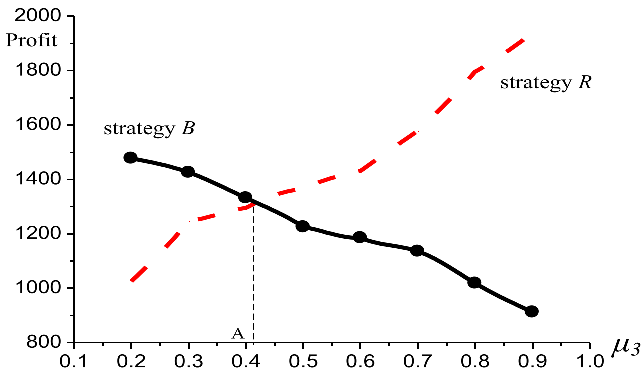


FIGURE 11. The impact of yield expectation on strategies in the extended model.

Observation 1:

When the yield's expectation (μ_3) is small, the main supplier can design a low price to compete with the rival under strategy B; if μ_3 is small, the manufacturer should use strategy B; otherwise, strategy R should be adopted.

The wholesale price and μ_3 under strategy B are showed in Fig.10. When μ_3 is small ($\mu_3 \leq A$), the manufacturer orders more components from the backup supplier to guarantee supply quantity. In order to compete with the backup supplier, the adaptive supplier would like to attract the manufacturer via a low wholesale price. For the manufacturer, the smaller purchase cost is a key incentive to choose more components from the main supplier. As a result, strategy B is more suitable when μ_3 is small. But when μ_3 is larger than the threshold ($\mu_3 > A$), the supply risk is mitigated and the manufacturer is more dependent on the main supplier. Hence, the main supplier should raise its own price under strategy B (Fig.10), which decreases the manufacturer's income simultaneously (Fig.11). Much replenishment costs can be saved if the manufacturer selects strategy R instead of B.

Then, the effects of the yield uncertainty level σ_3 on the manufacturer's strategies are different from the result in section IV.B.2).

Observation 2:

When the main supplier can determine the unit wholesale price, strategy *R* is not superior to strategy *B* when σ_3 is small.

Fig. 12 illustrates the effect of σ_3 on the main supplier's wholesale price under two strategies. It is interesting to find that, the wholesale price under strategy *R* is always higher. If strategy *R* is utilized, the main supplier becomes the unique replenishment channel. Naturally, the main supplier will decide a high wholesale price to earn more profits. Therefore, unlike the result in subsection IV.B.2), it is not beneficial to adopt strategy *R* even though the yield uncertainty level is small (as displayed in Fig.13). For strategy *B*, the main supplier's wholesale price is not high due to the competition with the backup supplier. Especially when σ_3 is large, the competitiveness of the main supplier is weakened by large yield uncertainty level. In order to attract more orders, it is sensible to cut down wholesale price for the main supplier under strategy *B*, which decreases the manufacturer's total cost. Thus, strategy *B* should be adopted when the yield uncertainty level is small.

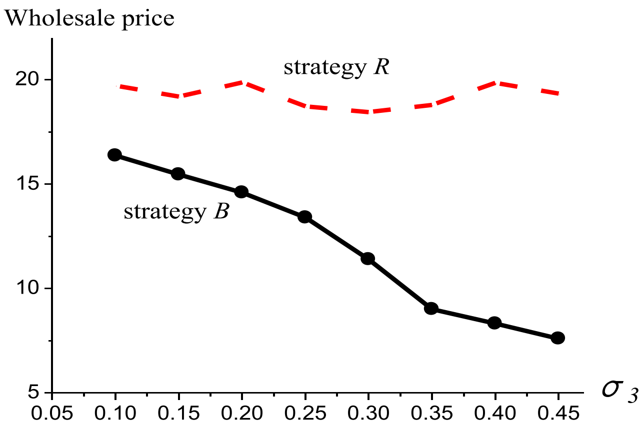


FIGURE 12. Yield uncertainty versus the main supplier's wholesale price under two strategies.

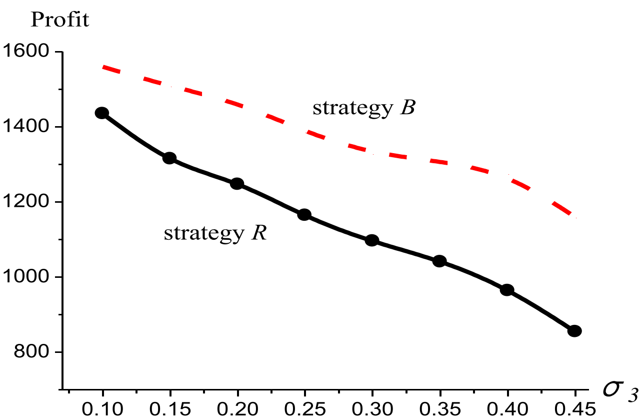


FIGURE 13. The impact of yield uncertainty on strategies in the extended model.

VI. CONCLUSION

We study an adaptive manufacturer's supply management strategies for mitigating uncertain yield risk from the supplier. The manufacturer has two supply management strategies: backup supplier (*B*) and reliability improvement (*R*). Under strategy *B* (*R*), the manufacturer dynamically adjusts price, order quantity, and reservation capacity (effort level) over multiple periods to maximize own profit. We study two strategies under the main supplier's exogenous wholesale price and endogenous wholesale price, respectively.

The impacts of consumer behavior and yield risk on the manufacturer's strategies are investigated. When the main supplier's unit wholesale price is exogenous, we show that: (i) if the number of instant consumers is sufficiently small, strategy *R* is the better choice; otherwise, it is suggested to adopt strategy *B*; (ii) if the expected quality sensitivity is not large, the manufacturer should use strategy *B*; or else the optimal strategy is *R*; and (iii) only when mean yield is large enough and yield uncertainty is small, the manufacturer prefers to use strategy *R*.

When the main supplier's wholesale price is endogenous, we observe that: (i) if the mean yield is small, the main supplier can decide a lower price to induce a higher order quantity from the manufacturer under strategy *B*; (ii) if the mean yield increases, the main supplier should raise the unit wholesale price; and (iii) strategy *R* should not be adopted even though the yield uncertainty level is small, which is inconsistent with the exogenous wholesale price setting.

There are several directions for future research. First, this paper only considers one manufacturer. This assumption could be relaxed to study a more complex model, where multiple manufacturers compete with each other. Second, the backup supplier is a non-adaptive agent. Therefore, it is worth examining the manufacturer's decisions when the backup supplier is endogenous. Third, coordination contract is an important tool to motivate collaboration for supply chain members. Hence, the issue of how to design a proper contract to maximize channel profit under yield risk should be further studied.

CONFLICT OF INTEREST

No conflict of interest

REFERENCES

[1] S. Deo and C. J. Corbett, "Cournot competition under yield uncertainty: The case of the US influenza vaccine market," *Manuf. Service Oper. Manage.*, vol. 11, no. 4, pp. 563–576, Oct. 2009.

[2] K. Chen and L. Yang, "Random yield and coordination mechanisms of a supply chain with emergency backup sourcing," *Int. J. Prod. Res.*, vol. 52, no. 16, pp. 4747–4767, 2014. doi: 10.1080/00207543.2014.886790.

[3] S. Y. Tang, H. Gurnani, and D. Gupta, "Managing disruptions in decentralized supply chains with endogenous supply process reliability," *Prod. Oper. Manage.*, vol. 23, no. 7, pp. 1198–1211, Jul. 2014.

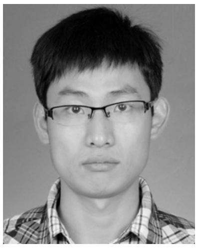
[4] D. R. Krause, "The antecedents of buying firms' efforts to improve suppliers," *J. Oper. Manage.*, vol. 17, no. 2, pp. 205–224, Jan. 1999.

[5] K. E. Train, *Discrete Choice Methods with Simulation*. New York, NY, USA: Cambridge Univ. Press, 2009.

[6] C. Bai and J. Sarkis, "Supplier development investment strategies: A game theoretic evaluation," *Ann. Oper. Res.*, vol. 240, no. 2, pp. 583–615, Nov. 2016.

[7] A. J. Schmitt, "Strategies for customer service level protection under multi-echelon supply chain disruption risk," *Transp. Res. B, Methodol.*, vol. 45, no. 8, pp. 1266–1283, Sep. 2011.

[8] J. M. Swaminathan, S. F. Smith, and N. M. Sadeh, "Modeling supply chain dynamics: A multiagent approach," *Decision Sci.*, vol. 29, no. 3, pp. 607–632, Jul. 1998.
[9] M. Henig and Y. Gerchak, "The structure of periodic review policies in the presence of random yield," *Oper. Res.*, vol. 38, no. 4, pp. 634–643, Aug. 1990.
[10] Y. Gerchak, "Order point/order quantity models with random yield," *Int. J. Prod. Econ.*, vol. 26, nos. 1–3, pp. 297–298, Feb. 1992.
[11] A. S. Erdem and S. Özekici, "Inventory models with random yield in a random environment," *Int. J. Prod. Econ.*, vol. 78, no. 3, pp. 239–253, Aug. 2002.
[12] R. Anupindi and R. Akella, "Diversification under supply uncertainty," *Manage. Sci.*, vol. 39, no. 8, pp. 944–963, Aug. 1993.
[13] B. Tomlin, "On the value of mitigation and contingency strategies for managing supply chain disruption risks," *Manage. Sci.*, vol. 52, no. 5, pp. 639–657, May 2006.
[14] V. Babich, A. N. Burnetas, and P. H. Ritchken, "Competition and diversification effects in supply chains with supplier default risk," *Manuf. Service Oper. Manage.*, vol. 9, no. 2, pp. 123–146, Apr. 2007.
[15] L. Dong and B. Tomlin, "Managing disruption risk: The interplay between operations and insurance," *Manage. Sci.*, vol. 58, no. 10, pp. 1898–1915, Jun. 2012.
[16] J. Cai, M. Zhong, J. Shang, and W. Huang, "Coordinating VMI supply chain under yield uncertainty: Option contract, subsidy contract, and replenishment tactic," *Int. J. Prod. Econ.*, vol. 185, pp. 196–210, Mar. 2017.
[17] G. Zhiyuan, Z. Meihua, and P. Hongjun, "Financing strategies for coal-electricity supply chain under yield uncertainty," *Int. J. Mining Sci. Technol.*, vol. 28, no. 2, pp. 353–358, Mar. 2018.
[18] S. Asian and X. Nie, "Coordination in supply chains with uncertain demand and disruption risks: Existence, analysis, and insights," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 44, no. 9, pp. 1139–1154, Sep. 2014.
[19] J. Chen, X. Zhao, and Y. Zhou, "A periodic-review inventory system with a capacitated backup supplier for mitigating supply disruptions," *Eur. J. Oper. Res.*, vol. 219, no. 2, pp. 312–323, Jun. 2012. doi: 10.1016/j.ejor.2011.12.031.
[20] J. Hou, A. Zeng, and L. Sun, "Backup sourcing with capacity reservation under uncertain disruption risk and minimum order quantity," *Comput. Ind. Eng.*, vol. 103, pp. 216–226, Jan. 2017.
[21] Z. Yin and C. Wang, "Strategic cooperation with a backup supplier for the mitigation of supply disruptions," *Int. J. Prod. Res.*, vol. 56, no. 12, pp. 4300–4312, 2018.
[22] A. Z. Zeng and Y. Xia, "Building a mutually beneficial partnership to ensure backup supply," *Omega*, vol. 52, pp. 77–91, Apr. 2015. doi: 10.1016/j.omega.2014.10.008.
[23] J. Hou and L. Zhao, "Backup agreements with penalty scheme under supply disruptions," *Int. J. Syst. Sci.*, vol. 43, no. 5, pp. 987–996, 2012.
[24] K. Chen and T. Xiao, "Production planning and backup sourcing strategy of a buyer-dominant supply chain with random yield and demand," *Int. J. Syst. Sci.*, vol. 46, no. 15, pp. 2799–2817, 2015. doi: 10.1080/00207721.2013.879234.
[25] M. Dada and R. Marcellus, "Process control with learning," *Oper. Res.*, vol. 42, no. 2, pp. 323–336, Apr. 1994.
[26] M. Khan, M. Y. Jaber, and A. R. Ahmad, "An integrated supply chain model with errors in quality inspection and learning in production," *Omega*, vol. 42, no. 1, pp. 16–24, Jan. 2014.
[27] M. S. Kim and B. Sarkar, "Multi-stage cleaner production process with quality improvement and lead time dependent ordering cost," *J. Cleaner Prod.*, vol. 144, pp. 572–590, Feb. 2017.
[28] R. L. Marcellus and M. Dada, "Interactive process quality improvement," *Manage. Sci.*, vol. 37, no. 11, pp. 1365–1376, Nov. 1991.
[29] E. L. Porteus, "Optimal lot sizing, process quality improvement and setup cost reduction," *Oper. Res.*, vol. 34, no. 1, pp. 137–144, Feb. 1986.
[30] H. Shu and X. Zhou, "An optimal policy for a single-vendor and a single-buyer integrated system with setup cost reduction and process-quality improvement," *Int. J. Syst. Sci.*, vol. 45, no. 5, pp. 1242–1252, Apr. 2014.
[31] M. R. Leenders and D. L. Blenkhorn, "Reverse marketing: The new buyer-supplier relationship," *J. Marketing*, vol. 53, no. 2, pp. 126–134, Apr. 1988.
[32] S. Liu, K. C. So, and F. Zhang, "Effect of supply reliability in a retail Setting with joint marketing and inventory decisions," *Manuf. Service Oper. Manage.*, vol. 12, no. 1, pp. 19–32, Feb. 2010.
[33] Y. Wang, Y. Xiao, and N. Yang, "Improving reliability of a shared supplier with competition and spillovers," *Eur. J. Oper. Res.*, vol. 236, no. 2, pp. 499–510, Jul. 2014.
[34] Y. Wang, W. Gilland, and B. Tomlin, "Mitigating supply risk: Dual sourcing or process improvement," *Manuf. Service Oper. Manage.*, vol. 12, no. 3, pp. 489–510, 2010.
[35] V. Gupta, B. He, and S. P. Sethi, "Contingent sourcing under supply disruption and competition," *Int. J. Prod. Res.*, vol. 53, no. 10, pp. 3006–3027, May 2015.
[36] L. Silbermayr and S. Minner, "Dual sourcing under disruption risk and cost improvement through learning," *Eur. J. Oper. Res.*, vol. 250, no. 1, pp. 226–238, Apr. 2016.
[37] H. A. Simon, "Theories of bounded rationality," *Decision Org.*, vol. 1, no. 1, pp. 161–176, 1972.
[38] Q. Q. Long, "A multi-methodological collaborative simulation for inter-organizational supply chain networks," *Knowl.-Based Syst.*, vol. 96, pp. 84–95, Mar. 2016.
[39] Q. Long, "Data-driven decision making for supply chain networks with agent-based computational experiment," *Knowledge-Based Syst.*, vol. 141, pp. 55–66, Feb. 2018.
[40] Z. Jiang, Y. Jin, E. Mingcheng, and Q. Li, "Distributed dynamic scheduling for cyber-physical production systems based on a multi-agent system," *IEEE Access*, vol. 6, pp. 1855–1869, Dec. 2018.
[41] Y. Zhou, J. J. Yang, and L. Y. Zheng, "Multi-agent based hyper-heuristics for multi-objective flexible job shop scheduling: A case study in an aero-engine blade manufacturing plant," *IEEE Access*, vol. 7, pp. 21147–21176, Feb. 2019.
[42] Z. He, S. Wang, and T. C. E. Cheng, "Competition and evolution in multi-product supply chains: An agent-based retailer model," *Int. J. Prod. Econ.*, vol. 146, no. 1, pp. 325–336, Nov. 2013.
[43] Y. He and X. Zhao, "Coordination in multi-echelon supply chain under supply and demand uncertainty," *Int. J. Prod. Econ.*, vol. 139, no. 1, pp. 106–115, Sep. 2012.
[44] B. Tomlin and Y. Wang, "On the value of mix flexibility and dual sourcing in unreliable newsvendor networks," *Manuf. Service Oper. Manage.*, vol. 7, no. 1, pp. 37–57, Jan. 2005.
[45] C. Li, "Sourcing for supplier effort and competition: Design of the supply base and pricing mechanism," *Manage. Sci.*, vol. 59, no. 6, pp. 1389–1406, Jan. 2013.
[46] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. Cambridge, MA, USA: MIT Press, 1998.
[47] C. J. C. H. Watkins, "Learning from delayed rewards," Ph.D. dissertation, King's College, Univ. Cambridge, Cambridge, U.K., 1989.
[48] U. W. Thonemann and J. R. Bradley, "The effect of product variety on supply-chain performance," *Eur. J. Oper. Res.*, vol. 143, pp. 548–569, Dec. 2002.
[49] C. Jiang and Z. Sheng, "Case-based reinforcement learning for dynamic inventory control in a multi-agent supply-chain system," *Expert Syst. Appl.*, vol. 36, no. 3, pp. 6520–6526, Apr. 2009.



HENG DU was born in Anhui, China, in 1990. He received the B.S. degree from Jiangxi Agricultural University, China, in 2012, and the M.S. degree from Southeast University, China, in 2015. He is currently pursuing the Ph.D. degree with Nanjing University, China. His research interests include supply chain management and multi-agent simulation.

YE JIANG was born in Jiangsu, China, in 1981. She received the B.S. degree from Nanjing Audit University, China, in 2005, and the M.S. degree from Xiamen University, China, in 2009. She is currently pursuing the Ph.D. degree with Nanjing University, China.

From 2009 to 2019, she was a Lecturer with the School of Business, Jiangsu Open University, Nanjing, China. Her research interests include supply chain and two-sided market.