*Proceedings of the 2022 International Symposium on Semiconductor Manufacturing Intelligence (ISMI2022)*

# Strategic Dynamic Pricing Optimization by Thompson Sampling and Stochastic Programming

Tran Hong Van Nguyen*, Chen-Fu Chien, Hsuan-An Kuo and Kang-Ting Ma

***Abstract*— Pricing Decision (PD) is considered as a strategic tool to generate profit, and dynamic pricing recently has been paid attention because it enables to deal with demand uncertainty and fast-changing market. Manufacturing companies are more likely to set their price following cost-based strategy without concerning on the price elasticity or customer preferences. In addition, to enhance yield management and maintain competitive advantages, capacity planning (CP) has been gotten more attention than PD in manufacturing companies. However, the lack of inter-communication between these two strategic activities in manufacturing companies could damage their overall business performance. Focusing on realistic needs, we propose an intelligent dynamic PD framework for jointly PD and CP by applying Thompson Sampling (TS) and Stochastic Programming (SP). First, to deal with demand-price function, this study applied TS to build prior distribution arbitrarily correlated for different prices and illustrate cross-elasticity for many products. Second, SP is employed to provide the optimal price policy under resource constraints. An empirical study is utilized to illustrate the proposed approach which showed the practical value on decision strategy for both PD and CP.**

***Keywords:*** manufacturing industry, Thompson sampling, Stochastic programming, Resource Constraints, dynamic pricing decision

## I. INTRODUCTION

Be evolved from airline industry, revenue management and dynamic pricing have gotten attention for both academia and industry in research topics and extensive application recently [1]. To acquire considerable gains in revenue and sustainable competitive advantages, companies have committed to maximize their profits over a time horizon under interdependent resources by changing products' prices as well as optimizing production planning. Along with advanced technology, Dynamic pricing strategy is increasingly set to become a vital factor in regulating inventory and lessening production pressure, thus considerately spreading out many industries such as retailers, hotel booking, food industry, and electric vehicle charging stations. Although little previous research has been done for manufacturers, manufacturing industry is potentially for revenue management methods because they have developed enormous sources of data availability and automated business process in customer relationship management and supply chain management.

In the presence of uncertain demand and partial information of demand curve, firms have been facing a series of decision problem to optimize future expected return based on past observation. Due to the lack of relationship information between price and demand, firms must implement some experiments and demand learning to gain knowledge, thus causing huge amount in expenditure. TS has received much attention in the past decade to solve the multi-armed bandit (MAB) problem by utilizing a Bayesian probabilistic approach [2]; so much work on the dynamic pricing strategy has been carried out by this algorithm [3], [4], [5]. The conventional approach helps to identify the price decision that maximizes expected profit with given posterior distribution over demand parameters, and then update the posterior distribution as a new response data. However, previous studies have mainly focused on theoretical methods, so there is still a room for applicable research providing managerial insights and numerical analytics.

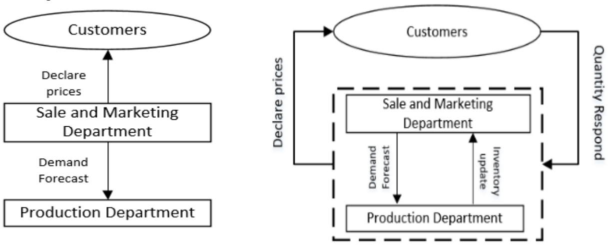


Figure 1(a). Current process Figure 1(b). Proposed process

A number of studies have found that classical operational problems, such as inventory management, yield management, and optimal CP, are generally decoupled from marketing and financial standpoint, especially PD [6]. In practical, the production department may construct the CP when receiving forecasted demand's value, then the sale and marketing department makes deal with customers based on cost-based pricing strategy which is sum of total manufacturing costs and a desired profit margin, as illustrated in Fig.1(a). The ignorance of demand feedback and nonalignment with available production capacity in this current pricing strategy might lead to unutilized capacity or unfulfilled demand [7]. Thus, along with CP problem in manufacturing firms, determination of

T.H.V. Nguyen, C.-F. Chien, and H.-A. Kuo are with the Department of Industrial Engineering and Engineering Management and the Intelligent Manufacturing and Circular Economy Research Center, National Tsing Hua University, Hsinchu 30013, Taiwan (e-mail: nguyentranhongvan@tdtu.edu.vn).

K.-T. Ma is with Department of Business Administration, National Dong Hwa, Taiwan and the Artificial Intelligence for Intelligent Manufacturing Systems (AIMS) Research Center, Ministry of Science and Technology, Taiwan.

products' price which has a strong influence on CP needs to be transformed by optimizing these two interrelated decisions as a whole. Therefore, this study proposes a closed loop process illustrated in Fig.1(b) which considers PD and CP altogether, and record demand quantity as feedback from customers to enhance the overall business performance.

Focusing on realistic needs in the manufacturing company, this study aims to propose a PD framework which aggregates TS and SP to solve the jointly PD and CP to empower Industry 3.5 [8]. The manufacturers enable to maximize their profitability over a finite time horizon by adjusting products' price settings aligning with production schedule. Particularly, considering of the relevant cost, stochastic demand and manufacturing resource constraints, the proposed framework could provide the suggestion prices of multiple products by online learning information from the demand curve with cross-elasticity concern, and then exploits this information to provide a near-optimal price decision strategy and even fulfill demand. TS samples demand from its posterior distribution and SP solves an optimization problem to incorporate inventory constraints and obtain optimal prices which illustrate high correlations of demand variations in the consecutive periods.

The reminder of this paper is constructed as follows. Section 2 mentions previous literature and applications of dynamic pricing strategy and MAB problems. Section 3 presents the problem definition and its mathematical model of Thompson Sampling- Stochastic Programming. Section 4 provides some numerical study to validate the proposed framework. Section 5 concludes this work with discussion of contributions and several future directions.

## II. LITERATURE REVIEW

Pricing Strategy and CP play crucial roles in generating overall profit for companies, thus there has been a surge of interest in researching of dynamic pricing with inventory effects. This research stream executes as a bridge to fill the gap between pricing problem and inventory management, and it has shown some promising applications. Kunreuther and Schrage [9] has proposed the joint static pricing and inventory problem with no capacity constraint which determine decisions for both pricing and ordering for one product under a deterministic demand curve that differs from period to period. In a study of van den Heuvel and Wagelmans [10], the author proposed a model with the consideration of uncapacitated economic lot-size to identify an optimal price decision as well as ordering number simultaneously with a deterministic demand function. Some researches do not changing price values because of potential adjustment costs (menu cost and managerial cost) and dissatisfaction of consumers [11]. Gayon and Dallery [12] has addressed the framework to provide the joint pricing strategies and optimal production, then showed that dynamic pricing is more likely superior than static pricing since the manufacturing line is not fully controlled. These previous researches assumed that the initial capacity level is fixed, thus several industries such as airline, restaurant, hotel, and music concert have been widely applied this method.

However, the decision of initial CP about the amount items of inventory for production or purchase is also a crucial decision to the firm, specially to manufacturing industry. PD and inventory management could be seen as a simultaneous optimization issue, thus leading to another stream which is named joint pricing and inventory-procurement problems [13], [14]. To deal with demand uncertainty, several robust and stochastic optimization approaches are applied to identify the joint pricing and inventory control problem [15]. Over a finite planning horizon, Chen and Hu [16] analyzed a joint pricing and inventory model to maximize the expected profit for a single product which an ordering quantity and product's price are implemented simultaneously. Since manufacturing company has to identify acquisition fare, selling price, manufacturing percentage, and remanufacturing rate, Mahmoudzadeh, et al. [17] has constructed a model for joint dynamic production and pricing problem that decisions is determined in each period confronting with uncertain demand and return. Based on a study of Chen and Lu [18], expected shadow price is employed to estimate price and two-stage scenario-based stochastic mixed integer programming model to illustrate the constraints of production capacity.

The MAB problem is seen as a classical research study which has attracted growing concern recently because of its wide applications in automated machine learning and practical domains [19]. In the family of algorithms to deal with balancing between immediate execution and experimentation, TS is a randomized Bayesian approach to solve online decision problems in which actions are sequentially implemented [20], and to explore the environment and explicitly address the trade-off problem between exploration and exploitation. TS randomly samples a parameter from the posterior distribution and derive a new set of actions by maximizing the expected reward for the next time step using the sampled parameters [21].

Several authors have experienced similar approach by using maximum likelihood estimators to derive exploration-exploitation problem in dynamic pricing [22], [23]. Besbes and Zeevi [24] have developed optimal price policies by dividing the horizon time into two phases where demand function is estimated at multiple prices in first phase, and revenue maximization problem is determined in second phase. Moradipari, et al. [4] has proposed a TS based algorithm for real-time pricing program under exogenously changing grid conditions. TS method has showed off some better practical outcomes [25], and outperform compared to other methods in same family such as the Greedy method and the Upper Confidence Bound algorithm [26], [27].

Because of the increase of intensive and competitive manufacturing market and huge pressure on transforming to smart manufacturing, adoption of digital transformation and information analysis are seen as effective and efficient solutions for decision-making. Thus, it is a research needed to address those interrelated elements simultaneously and provide applicable and managerial insights in realistic.

## III. PROPOSED METHODS

The proposed research framework in this study includes four distinct stages: (1) problem definition, (2) data preparation, (3) model construction, and (4) decision strategy, as illustrated in Fig.2.

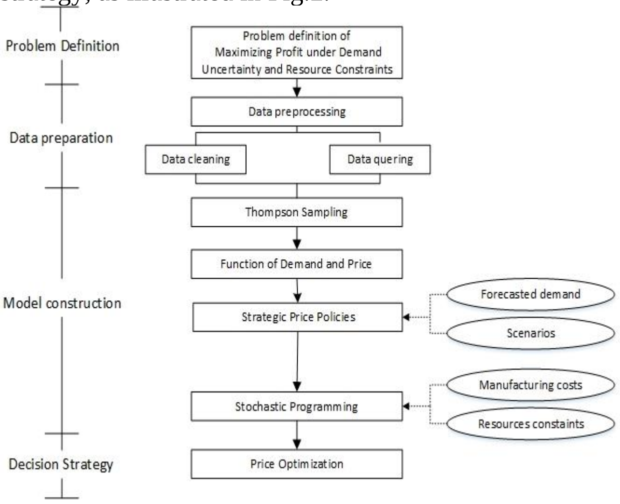


Figure 2. Proposed framework for price optimization

### *3.1. Problem definition*

In the era of high technology and digital transformation, companies are in the race to apply modern technology and data-driven method for revenue management, demand forecasting and production planning to reinforce their yield rate and productiveness. However, companies mostly determine optimal solutions of PD as an isolated part from production planning, so this lack of considering those inter-correlations may lead to sub-optimal values. An absence of uniformity of price setting in different products could break sale strategy of manufacturing companies and harm to overall effectiveness in which is shared interdependent sources such as machines, labors and raw materials. In addition, the setting of prices is mainly routine and intuitive state because of being merely dependent on managers' decision or experienced people. High price setting causes unattainability of the desired sales volume, while low price setting may result in extensive shortage of product availability, lose sales opportunities and an overall reduction in profitability. Industry 3.5 is defined as a hybrid strategy which is a forward step from Industry 3.0 to Industry 4.0 for disruptive innovations and needs for industrial migration. Along with domain knowledge, AI, advanced analytics, and big data technologies can be employed as novel solutions to empower digital decision. Thus, to move more analytical orientation for price setting, it is worthy to build a systematical dynamic PD in which considers several inter-correlated decisions. In addition, our study has focus on dynamic pricing and learning because historical demand data is digitally available that enables to lessen the efforts for acceptably accurate estimates of demand function.

### *3.2. Data Preprocessing*

To construct a dynamic pricing strategy, the raw data comprises price and quantity of many products. Once the automatic selling platform is recognized from salesman, it will be transformed and delivered to the local server of the company which stores the amount of quantity and selling price daily. Since the dataset is accumulated for one week, then it could upload and integrate the dataset from both marketing & sale departments and production department into the database for implementing the proposed framework. However, data manipulation, data entry errors, and data corruption could cause to noisy and missing values in the data. The stage of data preparation might be one of notoriously time-consuming and tedious stages, but this process is indispensable to assure data integrity and data quality for providing insightful and constructive decision-making. Therefore, data preparation such as data cleaning, data preprocessing, and transformation will be implemented prior to further analysis.

### *3.3. Model Construction and Decision Strategy*

#### *3.3.1. Description of model construction*

| **Indices:** | | **Parameters:** | |
|---|---|---|---|
| $j$ | Product index, $j = 1, 2, 3, \dots, n$ | $N$ | Total number of products |
| $i$ | Machine index, $i = 1, 2, 3, \dots, m$ | $M$ | Total number of machines |
| $k$ | Price index, $k = 1, 2, 3, \dots, k$ | $K$ | Total number of price point |
| $t$ | Time index, $t = 1, 2, 3, \dots, T$ | $S$ | Total number of scenarios |
| $s$ | Scenario index, $s = 1,2,3,\dots, s$ | $C$ | Total machine capacity |
| **Decision variables:** | | $r_{ji}$ | Process time of product $j$ at machine $i$ |
| $x^t_{jks}$ | Quantity of product $j$ to be produced under price policy l in period t and scenario s | $d_j$ | Expected demand of product $j$ |
| $l^t_j$ | $l^t_j$ = 1 if price policy l is chosen for product $j$ in period t; otherwise $l^t_j = 0$ | $p^t_j$ | Unit price of product j in period $t$ |
| | | $w_{iv}$ | Cross-price elasticity of product $j$ and product $v$ |
| | | $\delta_j$ | Price elasticity of product $j$ |
| | | $p_s$ | Probability of scenario $s$ |
| $D^t_{jk}$ | Demand of product $j$ with price $k$ in period $t$ | $c^t_j$ | Unit Cost of product j in period $t$ |

Fig.3 describes the time horizon for joint PD and production planning for manufacturing companies which following the proposed PD framework. At the beginning of period 1, firms start with the decision to amounts of producing product $j$ under a set amount of resources (labor, machine hours, raw materials) at unit manufacturing cost $c_j$ of a product $j$ per period. Thereafter, salesman from sale & marketing department declares the price of many products and accepts orders $q$ for product $j$ which will be fulfilled during the period, and any shortfall in capacity could be met through spot buying. After realizing the actual demand and inventory at the ending of period 1, firms will make the decision of production and price setting of products for period 2. This process keeps continuing for next period.

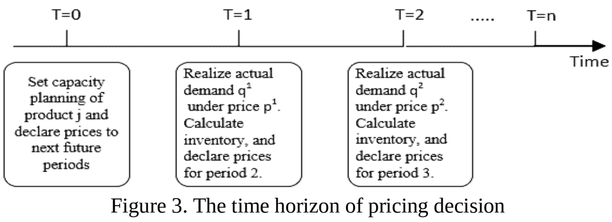


Figure 3. The time horizon of pricing decision

Related to the model construction, this study implies a model for the problem of joint dynamic pricing and inventory control in a monopoly setting. We apply constant-linear in terms of the price to review the market reaction of customers to the firm's product portfolio with cross-price elasticity. However, the real demand is uncertain, TS will be applied to simulate the demand distributions of each price suggested by balanced the exploration and exploitation. Action of agent in this case is a price policy chosen, and real demand quantity from customers is reward. Next, we match information of forecasted demand provided by the department of sale and marketing by minimum difference, we could find the respective price of for mean forecasted demand. Many different price policies should be taken into consideration of the PD support system (i.e. optimism, neural, and pessimism) depending on the styles of managers. In other words, some managers are eager to be the super star of year and choose more challenge with higher price, while others are more likely to be conservative to choose the lower price to meet their work requirement. To apply more realistic, SP is integrated into the TS to deal with resources constraints, and different scenarios of uncertain demand. In other words, the firm makes decision about pricing and numbers of products produced in the first stage, a recourse decision is then taken in the second stage based on the outcome of the decision in first stage.

### *3.3.2. Mathematical Formulation*

The demand model represents with discrete *k* price levels in which a price level is in accomplice with its intrinsic demand density function (pdf) described by several parameters. In a declarative manner and abstract the inference procedure, the generic Markov Chain Monte Carlo (MCMC) approach is applied to derive the model parameters because it is beneficial for a group of products with associated demand functions. In this study, the correlated parameters which are cross-elasticity of different demands will be identified from multivariate normal distribution, and probabilistic programming frameworks enables to determine and derive hierarchical models. To illustrate the existence of multivariate distribution for substitute products, and customize the model based on company constraints, the demand function for product *j* since chosen price *p* consists of two parts: first part is primary demand of this product *j* and second part is from substitutable demand of product *i*. Due to the effects of substitute products, since a product *j* is not available, its demand might be partially replaced by other products. Equation 1 presents the demand function of product *j* is which is influenced by its own price and prices of other substitute products.

$$D_j(P_j) = D_j + \sum\nolimits_{jn} w_{ij}\, D_{ij} \quad (1)$$

Where stands for $w_{ij}$ be the substitute rate of the demand of product *j* by another product *i*. With the constant- elasticity function, (1) could become (2):

$$D_j(P_j) = aP_j^{\gamma_j} + \sum bP_i^{\omega_{ij}} \quad (2)$$

Take the log-log from both sides, it becomes (3).

$$Log\, D_j(P_j) = Log\, a + \gamma_j LogP_j + \sum \omega_{ij} LogP_j \quad (3)$$

TS is applied ideas of MAB algorithm, and it uses the same parametric formula for demand function and proceeds to maximize revenue over a fixed horizon. The TS algorithm uses Gamma Posterior Distribution for each mean demand to implement updated rules for the distribution parameters in the probabilistic programming because it is a conjugate prior for the Poisson distribution. The prior distribution of mean demand is exponential with the cumulative distribution function $f(x) = e^{-x}$ which is equivalent to a Gamma distribution. The manufacturing company has offered price vector with the first period *t-1*. Initially, the posterior distribution of $D_{jk}^t$ follows Gamma distribution $W_{jk}(t-1)+1, N_k(t-1)+1)$, so we sample demand $D_{jk}^t$ of product *j* with the price level *k* at the time period *t*. The Model of Thompson Sampling- Stochastic Programing is summarized as follows:

*Step 1:* For each product, initialize a demand model by using Constant Cross - Elasticity function from equation (3)
*Step 2:* Sample demand from demand model for each product
Step 3: Identify the demand distribution with expected demand $D_{jk}^t$ by using absolute minimum difference of expected mean demand and forecasted demand provided $d_{jk}^t$.
*Step 4:* Split the range of demand distribution by quantiles
*Step 5:* Determine the decision for optimal price strategy for each product to have the maximum profit by applying SP

$$\text{Max} \sum_{s=1}^{S}\left(\sum_{k=1}^{K}\left(\sum_{j=1}^{N} D_{jk}^t\, P_j^t\, l_j^t\right)\times x_{jks}^t\right)\times p_s - c_j \times x_{jks}^t \quad (4)$$

$$\text{s.t: } x_{jks}^t \le D_{jk}^t \quad (5)$$

$$\sum l_j^t = 1 \quad (6)$$

$$l_j^t = \{0,\ 1\} \quad (7)$$

$$\sum_{1}^{M} r_{ji} \times x_{jks}^t \le C, m = 1,2,\ldots,M \quad (8)$$

$$p_s \in [0,1] \quad (9)$$

$$x_{jks}^t \ge 0,\ j = 1,2,\ldots,n$$

*Step 6:* Offer the price determined by SP and observe the true demand
*Step 7:* Update the posteriors of the selected price point for each product

Based on realistic requirements, a SP model in step 5 for production planning was formulated with following price strategy, and machine capacity constraints. The objective function (4) is to maximize profit. Constraints (5) express number of products which were chosen to produce must be less or equal to quantity sale of this period. Constraints (6) and (7) express price policy chosen, where $l_j^t$ is a binary variable which only single price policy is allowed to choose, and it is equivalent to one if Price Policy *h* is chosen for product *j* in period *t*, otherwise equals to zero. Constraint (8) illustrate for the constraint of processing time of machines, whereas Constraint (9) relates to the range of probability of each scenario from 0 to 1. This step is solved and the price vector $P(t) = p_{k*}$ for $k^*$ is the optimal price point chosen belong to price vector *[K]*; then then the actual demand $D_{jk}(t)$ is declared to manufacturing company at Step 6. Lastly, we update

$N_{k'}(t) \leftarrow N_{k'}(t-1)+1, W_{jk'}(t) \leftarrow W_{jk'}(t-1)+Dj(t)$ for all products $j \in [N]$.

## IV. NUMERICAL EXAMPLE

To validate the proposed approaches, a simulated data is applied to illustrate the proposed framework. Historical data is utilized to apply for the demand forecasting for the production planning, and material procurement. Thus, with the consideration of market changing, we develop a smart framework for dynamic pricing strategy by applying Thompson Sampling and Stochastic Programming to deal with demand uncertainty under resources constraint.

The historical data was collected to estimate parameters in proposed approaches. In following sections, this study presents the effectiveness of two approaches separately, and then introduces the implementation results with discussions. With the iteration of 500 sample time, Fig. 4 illustrates the demand distribution corresponding to each price level *k*, which the higher price acquires wider distribution and large variance. After those parameters of the constant-elasticity model are sampled, Fig. 5 exhibits the relationship between price and quantity, while the red dot represents historical points which is consistently suitable with the constant-elasticity demand function with the cross-elasticity concern.

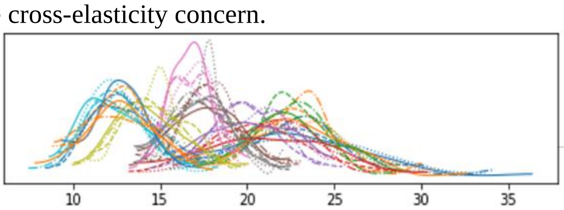

Figure 4. Demand distribution based on corresponding price

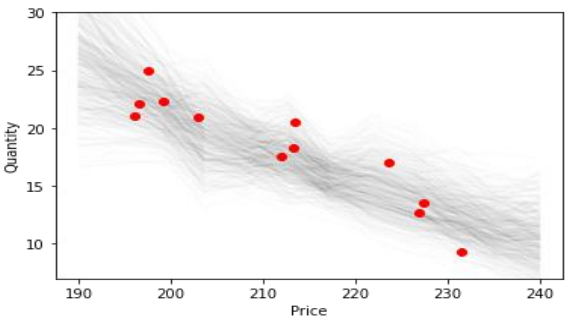


Figure 5. Relationship of quantity and price

Due to the uncertainty of demand, there are three forecasted demand value of product are provided by sale & marketing department including middle demand, high demand, and low demand for next periods. We applied the minimum different value between mean demand from TS algorithm and forecasted demand value to extract three corresponding demand distribution, shown in Fig. 6. Then, the interquartile range (IQR) is employed to illustrate the uncertain demand with equal probabilities of four scenarios as follows. In each scenario, the weighted arithmetic mean is calculated to emphasize frequency of some data points.

- Scenario 1 = (min value, Q1)
- Scenario 2 = (Q1, Q2)
- Scenario 3 = (Q2, Q3)
- Scenario 4 = (Q3, max value)

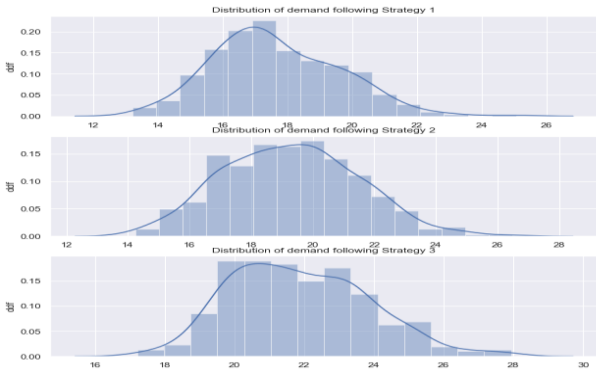


Figure 6. Distribution of demand following three price policies

A number of substitute products is launched in similar approach. In this study, three different management style including optimism, neural, and pessimism are considered corresponding to three forecasted demand of a product. Table 1 exhibits price policy description and forecasted demand. For instance, Price Policy P1 is called "Optimism" strategy which corresponds to high demand forecast might has lower price than Price Policy 2 that is called "Neural".

TABLE I. PRICE POLICY DESCRIPTION AND DEMAND SCENARIO

| *Price Policy* | *Price Strategies* | *Forecasted demand* |
|---|---|---|
| P1 | Optimism | High demand |
| P2 | Neural | Middle demand |
| P3 | Pessimism | Low demand |

Thus, a scenario-based stochastic programming is drawn from the output of TS algorithm, internal information of both forecasted demand value and capacity constraints, and domain knowledge. Gurobi software is utilized to solve the stochastic programming to provide the optimal solution of price strategy each week, as shown in Fig.7. The outcome from the proposed framework could provide a reference point to assist decision makers in PD and CP simultaneously.

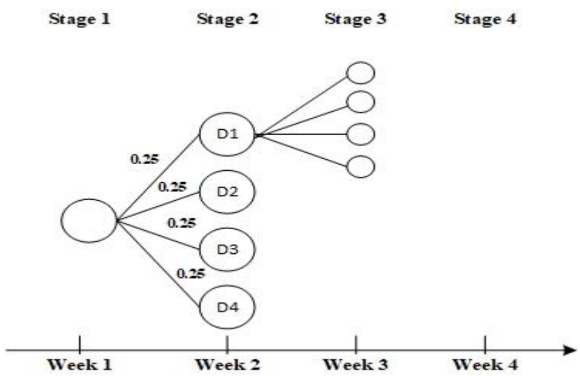


Figure 7. Multi-stage stochastic programming for PD

## V. CONCLUSION

The pricing strategy in manufacturing companies is mostly based on the cost-based pricing which describes the setting prices by adding a specific fixed percentage to cost of the goods and services. Thus, individual sellers in the marketing department are more likely to identify price setting by their personal experience, thus leading to an inefficient and unsystematic operational activity. In addition, there will be lacked the interaction between sell person man to other teams, which leads to the overall low effectiveness and efficiency of the company. Indeed, the price of a product enables to influence demand and revenue of other substitute products. If the market reaction is good for the product, cost-based pricing strategy is a type of “leaving money on the table” because it is failed to adjust to the environment.

To address realistic needs for jointly dynamic PD and CP, this study developed a dynamic PD framework integrating TS and SP to maximize profit under uncertain demand, and production constraints. It could solve for the trade-off exploration and exploitation problem, and active learning from updated values in which demand values and cost data are digitally available. Our study has several contributions as follows: (1) cross-price elasticity is included in demand function to illustrate of selling complement or substitute products; (2) scenarios analysis is mentioned for flexibility of management style; (3) stochastic factors is considered by taking uncertainty of demand into account; (4) dynamic pricing strategy is optimized with the presence of inventory constraints as a whole; (5) a digital decision support system was proposed to update and support companies decision-making activities about PD and CP. The results showed practical viability of the proposed method with realistic applications, and managerial insights.

This study has some limitations which could be improved in the future. First, to provide more comparable results, other optimization methods such as robust programming or dynamic programming can be integrated to TS. Second, other parameters including customer’s characteristics and market volatility can be integrated to demand function. Third, our study is limited by monopoly settings, examining the effect of competition and game theory could shed further light on researching of joint PD and CP.

## REFERENCES


[1] R. Klein, S. Koch, C. Steinhardt, and A. K. Strauss, "A review of revenue management: Recent generalizations and advances in industry applications," *European journal of operational research,* vol. 284, no. 2, pp. 397-412, 2020.

[2] S. Agrawal and N. Goyal, "Analysis of thompson sampling for the multi-armed bandit problem," in *Conference on learning theory*, 2012: JMLR Workshop and Conference Proceedings, pp. 39.1-39.26.

[3] K. J. Ferreira, D. Simchi-Levi, and H. Wang, "Online network revenue management using thompson sampling," *Operations research,* vol. 66, no. 6, pp. 1586-1602, 2018.

[4] A. Moradipari, C. Silva, and M. Alizadeh, "Learning to dynamically price electricity demand based on multi-armed bandits," in *2018 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, 2018: IEEE, pp. 917-921.

[5] M.-h. Oh and G. Iyengar, "Thompson sampling for multinomial logit contextual bandits," *Advances in Neural Information Processing Systems,* vol. 32, 2019.

[6] G. Bitran and R. Caldentey, "An overview of pricing models for revenue management," *Manufacturing & Service Operations Management,* vol. 5, no. 3, pp. 203-229, 2003.

[7] S. Deng and C. A. Yano, "Joint production and pricing decisions with setup costs and capacity constraints," *Management Science,* vol. 52, no. 5, pp. 741-756, 2006.

[8] C.-F. Chien, T.-y. Hong, and H.-Z. Guo, "A conceptual framework for “Industry 3.5” to empower intelligent manufacturing and case studies," *Procedia Manufacturing,* vol. 11, pp. 2009-2017, 2017.

[9] H. Kunreuther and L. Schrage, "Joint Pricing and Inventory Decisions for Constant Priced Items," *Management Science,* vol. 19, no. 7, pp. 732-738, 1973, doi: 10.1287/mnsc.19.7.732.

[10] W. van den Heuvel and A. P. M. Wagelmans, "A polynomial time algorithm for a deterministic joint pricing and inventory model," *European Journal of Operational Research,* vol. 170, no. 2, pp. 463-480, 2006.

[11] G. P. Cachon and P. Feldman, "Dynamic versus static pricing in the presence of strategic consumers," in *Working paper, University of Pennsylvania*, ed. Philadelphia, 2010.

[12] J.-P. Gayon and Y. Dallery, "Dynamic vs static pricing in a make-to-stock queue with partially controlled production," *OR Spectrum,* journal article vol. 29, no. 2, pp. 193-205, April 01 2007, doi: 10.1007/s00291-006-0049-0.

[13] W. Elmaghraby and P. Keskinocak, "Dynamic pricing in the presence of inventory considerations: Research overview, current practices, and future directions," *Management science,* vol. 49, no. 10, pp. 1287-1309, 2003.

[14] Ö. Özer and R. Phillips, *The Oxford handbook of pricing management*. OUP Oxford, 2012.

[15] E. Adida and G. Perakis, "Dynamic pricing and inventory control: robust vs. stochastic uncertainty models—a computational study," *Ann Oper Res,* journal article vol. 181, no. 1, pp. 125-157, December 01 2010.

[16] X. Chen and P. Hu, "Joint pricing and inventory management with deterministic demand and costly price adjustment," *Operations Research Letters,* vol. 40, no. 5, pp. 385-389, 2012.

[17] M. Mahmoudzadeh, S. J. Sadjadi, and S. Mansour, "Robust optimal dynamic production/pricing policies in a closed-loop system," *Applied Mathematical Modelling,* vol. 37, no. 16, pp. 8141-8161, 2013.

[18] T.-L. Chen and H.-C. Lu, "Stochastic multi-site capacity planning of TFT-LCD manufacturing using expected shadow-price based decomposition," *Applied Mathematical Modelling,* vol. 36, no. 12, pp. 5901-5919, 2012.

[19] D. Bouneffouf, I. Rish, and C. Aggarwal, "Survey on applications of multi-armed and contextual bandits," in *2020 IEEE Congress on Evolutionary Computation (CEC)*, 2020: IEEE, pp. 1-8.

[20] W. R. Thompson, "On the likelihood that one unknown probability exceeds another in view of the evidence of two samples," *Biometrika,* vol. 25, no. 3-4, pp. 285-294, 1933.

[21] D. J. Russo, B. Van Roy, A. Kazerouni, I. Osband, and Z. Wen, "A tutorial on thompson sampling," *Foundations and Trends® in Machine Learning,* vol. 11, no. 1, pp. 1-96, 2018.

[22] J. Broder and P. Rusmevichientong, "Dynamic pricing under a general parametric choice model," *Operations Research,* vol. 60, no. 4, pp. 965-980, 2012.

[23] A. V. den Boer and B. Zwart, "Simultaneously learning and optimizing using controlled variance pricing," *Management science,* vol. 60, no. 3, pp. 770-783, 2014.

[24] O. Besbes and A. Zeevi, "Dynamic pricing without knowing the demand function: Risk bounds and near-optimal algorithms," *Operations Research,* vol. 57, no. 6, pp. 1407-1420, 2009.

[25] O. Chapelle and L. Li, "An empirical evaluation of thompson sampling," *Advances in neural information processing systems,* vol. 24, 2011.

[26] N. Gupta, O.-C. Granmo, and A. Agrawala, "Thompson sampling for dynamic multi-armed bandits," in *2011 10th International Conference on Machine Learning and Applications and Workshops*, 2011, vol. 1: IEEE, pp. 484-489.

[27] E. Kaufmann, N. Korda, and R. Munos, "Thompson sampling: An asymptotically optimal finite-time analysis," in *International conference on algorithmic learning theory*, 2012: Springer, pp. 199-213.