

Research Article

A Bilevel Stochastic Dynamic Programming Model to Assess the Value of Information on Actual Food Quality at Wholesale Markets

Xiangyu Hou,¹ Rene Haijema,² and Dacheng Liu¹

¹Department of Industrial Engineering, Tsinghua University, Beijing 100084, China

²Operations Research and Logistics, Wageningen University, Hollandseweg 1, 6706 KN Wageningen, Netherlands

Correspondence should be addressed to Xiangyu Hou; houxu0727@gmail.com

Received 13 May 2017; Accepted 13 September 2017; Published 31 October 2017

Academic Editor: Mauro Gaggero

Copyright © 2017 Xiangyu Hou et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In the fresh produce wholesale market, the market price is determined by the total demand and supply. The price is stochastic, and either wholesaler or retailer has few influence on it. In the wholesaler's inventory decision, the price's uncertainty plays an important role as well as the uncertainty from the demand side: the wholesaler makes his decision based on the retailer's ordering, which is influenced by the stochastic market price and the distribution of the consumer's demand. In addition, when at the wholesale stage, the products show a similar quality of similar appearance. With more efforts being input, the wholesaler could detect and record more additional information than that reflected from the appearance. Based on this, he can classify the quality into different levels. No experience shows how the wholesaler could use the underlying quality information and how much this information could improve his profit. To describe and explore this problem, a bilevel dynamic programming approach is employed. We evaluate different strategies of using the underlying information, show the features of the optimal policy, develop heuristics, and discuss the influence of factors such as quality and market price. We also develop the managerial principles for the practical use.

1. Introduction

In China, the wholesale market plays a key role in food supply chains. The market is in the center of the chain: it matches supplies by wholesalers and demands by retailers in perfect competition, and thus prices are set. The price in the perfect market is essentially determined by the supply and demand [1]. The government occasionally interferes by controlling the supply and demand [2], when the price tends to get below or above an acceptable range. The price stochasticity is mainly caused by the uncertain factors in the supply side such as weather, yield of production. For those products without seasonal characteristics, the price change is usually gradual and shows mean-reverting feature due to the supply and demand's response to the price fluctuation as well as the intervention of government. The fluctuating market price has effects on both the wholesaler's and the retailer's inventory decisions. As a result, their inventory decisions rely on not

only the inventory position and the demand, but also the market price.

In a market with perfect competition, there is full information exchange making the products' quality look similar, but the products are usually sold differently at the retail stage. At the wholesale stage, for the products with relatively long shelf life, their quality converges to a certain range in which there is no visible difference in appearance. However, the underlying quality difference becomes apparent over time and evolves to visible quality difference at the retail stage. Besides, the consumers' more careful selecting for their own use also intensifies this phenomenon. Nowadays, except observing the appearance of product, the wholesaler can obtain the underlying information by recording the whole process from picking to consuming or by monitoring relevant parameters like the temperature, the gas composition of the environment, and so on. If the wholesaler obtains this information, information asymmetry occurs between the

wholesaler and the retailer, and it is possible to improve the profit by taking advantage of it. How to make use of it and how much the improvement could be are still not clear for the practitioners.

We also notice some other features. (1) The wholesaler brings products to the market from the producers hundreds or thousands of kilometers away, while the retailer is usually near the market [1]. The difference in distance causes different lead time. As a result, the wholesaler has to make ordering decision in advance and without knowing the influence of the market price exactly, while the retailer knows the current day's price and makes his decision without any uncertainty on price. (2) With the underlying difference of quality not shown visually, the wholesaler can dispose the products more easily and can get a higher salvage value comparing to when the products are disposed by the retailer.

In this paper, the decision of the wholesalers is modelled as a stochastic dynamic programming (SDP) model. With the retailer's decision process embedded in the model, it becomes a bilevel SDP. Using this model, we assess the value of the underlying quality information obtained by the wholesaler. We also discuss the optimal inventory policies for the wholesaler and experiment on different scenarios to show how to adjust the policy and how to use quality information under different conditions.

The paper will be organized as follows: in Section 2 we review the related literature and then come to the detailed problem description in Section 3, where we formulate the problem as a bilevel Markov decision process model. In Section 4, we explore how additional information on product quality can be used in a realistic base case. By simulation we explore the structure of the optimal inventory policy, which results in practical ordering policies. In Section 5, the approaches are tested for varying problem settings. In the last section, the findings and contribution of this paper are summarized, and limitations and directions for future research are discussed.

2. Literature

For modelling operational problems in the wholesale market, it is essential to include price uncertainty, as argued in Garnaut et al. [3], Li [2], and Watson [1]. The problem of combining stochastic spot market price with inventory control is widely studied. Kalyon [4] uses a Markov process to describe the price fluctuation. The market price has a mean-reverting feature, which is usually modelled as an Ornstein-Uhlenbeck (OU) process, for instance, Berling and Martínez-de-Albéniz [5] and Chen et al. [6]. When modelling the problem in discrete time as a Markov decision process, we depict the stochastic price based on a discrete version of an OU process, that is, an Ehrenfest chain. Besides, it should be noted that although we consider one wholesaler and one retailer in the market, their transaction price follows the stochastic market price in reality. Hence, we introduce this feature of market price into our inventory model.

Another important aspect is the quality. According to Grunert [7], there are many ways to define food *quality* as an indicator for the perceived freshness. In his framework

of analyzing consumer's perception on food quality, time (quality perception before/after purchase) is one of the two major dimensions together with the inference-making (how consumers infer quality from a variety of signals or cues). As we introduced, with the product decaying, the quality perception after purchase could have increasing difference in appearance comparing to the quality perception before purchase. To reduce its influence, the wholesaler could obtain more accurate quality information before purchase by monitoring and recording other relevant information with new technology. This kind of new technology's influence is supported by many reports. For instance, Bertolini et al. [8] report the impact of RFID on managing perishable inventory. We want to study the influence of additional quality information, which causes information asymmetry between the wholesaler and retailer. The information asymmetry and the signaling mechanism on the unobservable quality are reviewed by Kirmani and Rao [9], and Hobbs [10] analyzes and classifies the traceability system in resolving information asymmetry, but the problem for the fresh quality is not fully studied without a clear description on the concept or the mechanism. To describe the mechanism and influence in operational level, we follow Ferguson and Ketzenberg [11], in which the demand is connected with quality and the quality is expressed by inventory state. The after-purchase quality perception's influence is accessible for the wholesaler, while traditionally the wholesaler and retailer can only observe the before-purchase quality (in the perfect market this quality converges to the same value) and estimate the influence on the retail stage. The information asymmetry caused by extra effort in monitoring quality in the market is different from the studies before, and the way we study it in the operational level is also different.

In a perishable inventory problem, the wholesaler's ordering decision affects the availability and quality of product. But with the quality's influence not clearly shown at the wholesale stage, and also with the stochastic market price, the retailer might have his own decision problem. To the best of our knowledge, no similar research includes this kind of retailer's decision into the wholesaler's ordering decision. We would like to evaluate how much improvement can be achieved by including the considerations above into the wholesaler's inventory model. Following the Markov decision process framework of Haijema [12], the wholesaler's replenishment and disposal decisions could be well modelled. The quality information works on both the wholesaler and the retailer. Also, their goals might diverge in the supply chain. We could include the quality information and model this divergence by extending the MDP to a bilevel structure. This method is reviewed in Colson et al. [13] and is applied to different problems. For instance, in van Dijk et al. [14], a bilevel stochastic dynamic programming (SDP) model is used to analyze the fishery policy. With similar structure in our problem, we could replace fishery policy in the first level with the wholesaler's inventory problem similar to Haijema [12] and replace the fishermen's decision model in the second level by a Newsvendor model for the retailer, who is close to the market and could make single period decisions. Besides, the information asymmetry can be defined in the model such that

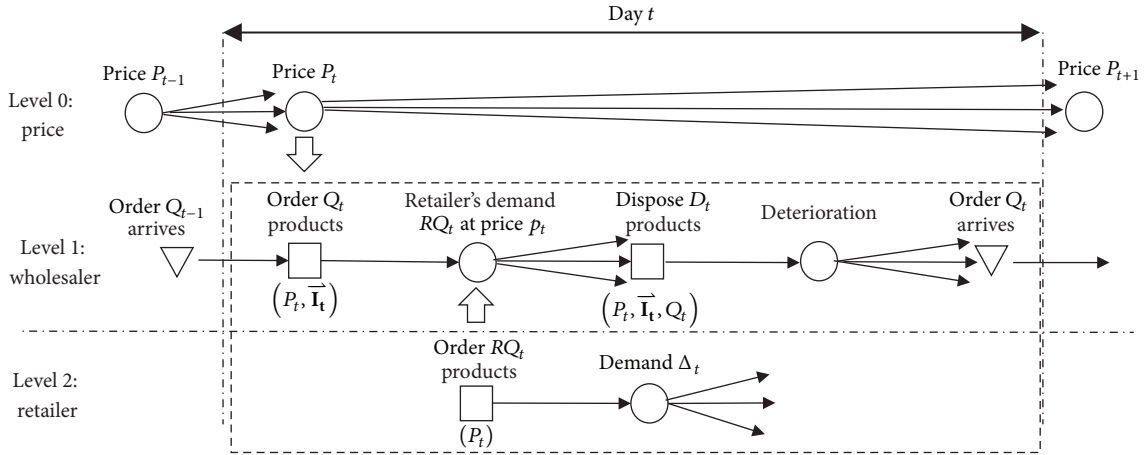


FIGURE 1: The discrete time model of the system.

the inventory state is unknown to the retailer and the offering sequence is fully controlled by the wholesaler to be a pure FIFO.

3. Bilevel Stochastic Dynamic Programming Model

In this section, we first describe the events in the market in detail. Then, we translate the features of market into model assumptions with mathematical language and formulate the problem as a bilevel model. We also discuss the solving procedure in this section.

3.1. Discrete Time Model. The decision-making of the wholesaler and the retailer is modelled as a bilevel optimization problem. The processes in both levels as well as the system state change are shown in Figure 1. The wholesaler's inventory state change and his decision are in Level 1, and the retailer's decision is in Level 2. The market price's change is exogenous; hence we isolate it from the bilevel model as Level 0.

(i) Level 0: The Market Price. The market price P_t influences both the wholesaler's and the retailer's decisions, but none of their decisions can determine the market price. The market price reflects the total supply and demand of all wholesalers and retailers, so it is set to be exogenous. The total supply and demand are influenced by uncertain factors, so the price is stochastic. The adjustment of the total supply and demand is based on the situation in the previous day and approaching to a balanced result, so the price shows mean-reverting feature.

(ii) Level 1: The Wholesaler's Inventory. At Level 1, day t begins after the order Q_{t-1} arrives. Then the wholesaler places a new order of Q_t at his supplier which will arrive at the end of day t and will be sold from day $t + 1$ on. After ordering the wholesaler sells the products in stock with FIFO/LIFO/mixed sequence at an uncertain price of that day. The demand of the retailer depends on the price and the quality of the products in stock at the wholesaler and is the result of an optimization at Level 2.

Then the wholesaler disposes some products to the secondary market. If the old products are disposed, product quality in the second day is promoted (the after-purchase quality is improved, but the before-purchase quality shows no difference in product's appearance). After these actions, the product deteriorates randomly, the products ordered at the start of that day are added to stock, and the next day comes.

(iii) Level 2: The Retailer's Decision. As introduced, the retailer is geographically close to the market and he goes to the market every day. The consumers at the retailer require high quality; therefore any leftovers are sold to a secondary market at a salvage value. Hence we assume retailer makes single period decisions. The retailer bases his purchasing decision on the market price of the current day (he knows the exact value) and the stochastic consumer's demand (he only knows the distribution). The underlying quality becomes visually apparent at the retail stage, so the quality's influence on the demand's distribution is also included. The quality of retailer's products together with the stock state is influenced by the wholesaler's decision at Level 1.

3.2. Assumptions and Model. Then we establish some assumptions to describe the process in the market and make a clear definition of the problem.

3.2.1. Global Setting: The Expression on Inventory and Deterioration. We consider a product with 2 underlying quality levels, denoted as $k = 1$ and $k = 2$ separately. The inventory state in each quality level is underlying information (or after-purchase information as discussed). The wholesaler could discover it by monitoring other relevant information. The inventory could be expressed as a vector $\vec{I}_t = (I_{1,t}, I_{2,t})$, where $I_{1,t}, I_{2,t}$ represent the number of low/high-quality products separately (i.e., $k = 1$ and $k = 2$). The vector stochastically shifts every day (representing the deterioration), and $\vec{I}_{t+1} = (I_{1,t+1}, I_{2,t+1})$.

The quantity and quality are the key features we are monitoring. The quantity is expressed by the vector above,

and the quality is defined by the weighted quality over different quality levels, that is, $\theta_t = \sum_{j=1}^2 (j/2) I_{j,t} / \sum_{j=1}^2 I_{j,t}$ when products with different qualities are mixed together to sell. This assumption follows Ferguson and Ketzenberg [11]. The wholesaler calculates θ_t from \vec{I}_t , and the retailer may also know it only if the wholesaler let him know the inventory state, or the retailer could estimate it as $\hat{\theta}_t$.

The product deteriorates stochastically. We assume the amount of decayed products to be random. It satisfies a binomial distribution with the decay probability $\vec{\beta}$. The probability that X products decay is $P(\xi_k = X) = \binom{I_{k,t}}{X} \beta_k^X (1 - \beta_k)^{I_{k,t}-X}$ (so, for any unit of product, the expected shelf life is $\sum_{k=1}^2 1/\beta_k$). And the inventory transition is as follows:

$$I_{k,t+1} = \begin{cases} I_{k,t} - \xi_k + \xi_{k+1}, & k = 1, \\ I_{k,t} - \xi_k, & k = 2. \end{cases} \quad (1)$$

3.2.2. Level 0: Market Price. Before the two levels for the wholesaler and the retailer, the global settings (such as the property of the market, or of the product) should be considered first. In this part, the most important variable is the market price, which works on both the wholesaler and the retailer. We model the process of price change at Level 0.

As discussed, the market prices in two successive days are similar, and the price has mean-reverting feature. We model it in a discrete way similar to an Ehrenfest chain. Assume the average market price is p_0 , and it could be $p_0 + x\gamma_p$ with $x \in \{-m, -m+1, \dots, -1, 0, 1, \dots, m-1, m\}$ and $\gamma_p > 0$ is the step size. The transition probability of price change satisfies $\Pr(P_0 + x\gamma_p, P_0 + x\gamma_p) \stackrel{\text{def}}{=} \Pr(x, x) = \phi$, $\Pr(P_0 + x\gamma_p, P_0 + (x+1)\gamma_p) \stackrel{\text{def}}{=} \Pr(x, x+1) = (1 - \phi)((m - x)/2m)$, and $\Pr(P_0 + x\gamma_p, P_0 + (x-1)\gamma_p) \stackrel{\text{def}}{=} \Pr(x, x-1) = (1 - \phi)((m + x)/2m)$, where ϕ is the probability that the price remains the same. It means that the price can only remain the same or transit to the adjacent states in two successive days.

3.2.3. Level 2: The Retailer. The consumers are sensitive to the quality difference shown at the retail stage. To start every day with the freshest products available, the retailer disposes or sells its remaining stock at the end of the day at a salvage value. The retailer visits the market every day, which provides him with the convenience to operate in a daily cycle. We assume the retailer applies a Newsvendor model. The stochastic consumer's demand relates to the quality, so the retailer's order varies depending on the quality as well as the market price. We assume the consumer's demand Δ_t follows a Poisson distribution with parameter λ_t . $\lambda_t = \phi_t A$ and A is a scale parameter, and ϕ_t is a coefficient estimating the quality's influence of the products bought from the wholesaler. It relates to the quality; that is,

$$\phi_t = (1 - \alpha + \alpha\theta_t), \quad (2)$$

where α measures the degree of quality's influence, and it is a constant parameter in $[0, 1]$. Usually θ_t is the underlying

information which the retailer does not know. It is estimated to an empirical value $\hat{\theta}_t$.

In reality, the consumer's perception on quality should be based on the products the retailer keeps, which are determined by the retailer's ordering RQ_t . But when the retailer makes his ordering decision, he cannot know the quality until he gets the products. So he can only estimate an empirical value $\hat{\theta}_t$ or use the information shared by the wholesaler (i.e., $\theta_t = \sum_{j=1}^2 (j/2) I_{j,t} / \sum_{j=1}^2 I_{j,t}$). Using the wholesaler's information is also an approximation for the retailer to make ordering decision, but it does reflect the influence of the wholesaler's decision and avoid the iteration in the solving procedure (if this approximation is not applied, i.e., the retailer uses the quality of the products he keeps, his decision will influence the quality and the quality will influence his decision, which requires iteration).

(i) *States.* The retailer makes his decision based on the market price and quality (estimated from the inventory state). As assumed, the state space of the price P_t is $\{P_0 - m\gamma_p, P_0 - (m-1)\gamma_p, \dots, P_0, \dots, P_0 + m\gamma_p\}$. We first consider the retailer knows the wholesaler's inventory state; hence he estimates the quality from the wholesaler's inventory state \vec{I}_t , which satisfies $\{(I_{1,t}, I_{2,t}) \mid |\vec{I}_t| = I_{1,t} + I_{2,t} \leq \Omega\}$ where Ω is the wholesaler's capacity.

(ii) *Actions.* The retailer orders RQ_t from the wholesaler, and he can order $|\vec{I}_t|$ at most. So the action space is $\{0, 1, \dots, |\vec{I}_t|\}$.

(iii) *State Transitions.* Since the retailer makes single period decision, his action does not affect the state of the next period. No state transitions need to be considered in the second level.

(iv) *Contribution.* The daily revenue of the retailer includes two parts, the revenue selling to the consumer and the salvage value. We only consider the purchasing cost. So the daily profit is as follows:

$$\begin{aligned} \max_{0 \leq RQ_t \leq |\vec{I}_t|} \Pi_2(P_t, \vec{I}_t) &= \mathbb{E} \max_{0 \leq RQ_t \leq |\vec{I}_t|} \{RP \min\{RQ_t, \Delta_t\} \\ &+ s_R(RQ_t - \Delta_t)^+ - P_t RQ_t\}. \end{aligned} \quad (3)$$

3.2.4. Level 1: The Wholesaler. As Haijema [12], the ordering and disposing are the wholesaler's decisions. But when offering the products to the retailer, a FIFO or a LIFO sequence could be applied depending on the wholesaler and retailer's perception on quality. The retailer's ordering quantity is the outcome of the decision model in Level 2 rather than simply from a random variable.

To keep the discussion clear, we assume that when the wholesaler knows the underlying quality and keeps it private, he provides product to the retailer with a FIFO sequence. While he shares the information with the retailer, a LIFO sequence is applied. If both the wholesaler and retailer do not know the underlying quality information, the sequence is a mix of LIFO and FIFO which is uncontrollable and the proportion is unknown.

(i) *States*. The retailer shares the same states as the wholesaler. So the wholesaler's state at Level 1 is the same as that at Level 2 (P_t, \vec{I}_t). The same is the state space.

(ii) *Actions*. The wholesaler orders Q_t from the producer and disposes D_t to the secondary market. $D_t \leq |\vec{I}'_t|$, where $|\vec{I}'_t|$ is the inventory after sales. $Q_t \leq \Omega - |\vec{I}'_t| + D_t$.

(iii) *State Transitions*. The state transition is split into two parts, as the transition relates to price and to stock levels.

The price state transition is

$$P_{t+1} = \begin{cases} P_t + \gamma_p, & \text{with probability } (1 - \phi) \frac{m - (P_t - P_0)/\gamma_p}{2m}, \\ P_t - \gamma_p, & \text{with probability } (1 - \phi) \frac{m + (P_t - P_0)/\gamma_p}{2m}, \\ P_t, & \text{with probability } \phi. \end{cases} \quad (4)$$

The inventory state transition is presented by first subtracting the sales that are met in either FIFO order or in LIFO order (depending on whether information on product quality is shared) and next subtracting any disposed and deteriorated products and, finally, adding new products. In case of information asymmetry, sales are met by issuing the oldest items (FIFO):

$$I'_{k,t} = \begin{cases} (I_{1,t} - RQ_t)^+, & k = 1, \\ (I_{2,t} - (RQ_t - I_{1,t})^+)^+, & k = 2. \end{cases} \quad (5)$$

When the quality information is shared with the retailer, the youngest/freshest products are issued first (LIFO):

$$I'_{k,t} = \begin{cases} (I_{2,t} - RQ_t)^+, & k = 2, \\ (I_{1,t} - (RQ_t - I_{2,t})^+)^+, & k = 1. \end{cases} \quad (6)$$

After the disposal decision (dispose the old product first),

$$I''_{k,t} = \begin{cases} (I'_{1,t} - D_t)^+, & k = 1, \\ (I'_{2,t} - (D_t - I'_{1,t})^+)^+, & k = 2. \end{cases} \quad (7)$$

After the deterioration,

$$I'''_{k,t} = \begin{cases} I''_{1,t} - \xi_1 + \xi_2, & k = 1, \\ I''_{2,t} - \xi_2, & k = 2. \end{cases} \quad (8)$$

After the order arrived,

$$I_{k,t+1} = \begin{cases} Q_t + I'''_{2,t}, & k = 2, \\ I'''_{1,t}, & k = 1. \end{cases} \quad (9)$$

(iii) *Contribution*. The retailer orders RQ_t . The purchasing cost from the supplier is c_0 . The wholesaler has to pay an

additional ordering cost K for each order. The part D_t selling to the secondary market is getting a salvage value s_w . The shortage cost (penalty) for the wholesaler is c_p . The daily holding cost per unit is c_h . So the wholesaler's daily profit is given by

$$\max_{Q_t, D_t} \Pi_1(P_t, \vec{I}_t) = \mathbb{E} \max_{Q_t, D_t} \left\{ P_t \min \{ RQ_t, |\vec{I}_t| \} - c_0 Q_t - K \delta(Q_t) + s_w D_t - c_p (RQ_t - |\vec{I}_t|)^+ - c_h |\vec{I}_t| \right\}. \quad (10)$$

3.3. Solution Procedure. At Level 2, the retailer's optimization decision satisfies a Newsvendor model, whose conclusion can be easily applied to the subproblem.

At Level 1, the retailer's decision process at Level 2 is called in every state transition. With this decision process embedded and the price's fluctuation included, the state transition matrix has a multichain structure. We apply a policy iteration algorithm to solve the model with average reward criterion, which is clearly introduced in Puterman [15]. The optimality and convergence are also given in Puterman [15].

When the quality of the wholesaler's product is unknown to the retailer (either the wholesaler also does not know it, or the wholesaler chooses to keep the information private), the retailer estimates the quality and makes ordering decision first, and then the wholesaler determines his policy based on it. Then the retailer adjusts his estimation and the wholesaler updates his policy iteratively. To be able to solve the problem numerically, we consider two quality classes. This is not uncommon in practice, although in theory one could think of more quality classes. The policy solved from SDP leads to a quality perception that is tracked using the quality parameter θ_t . To avoid the iteration on θ_t , we consider the following six values (0.5, 0.6, 0.7, 0.8, 0.9, 1.0) as the estimation of θ_t . For each value, we solve the SDP and simulate the resulting policy, as if θ_t is constant. The simulation reveals the real fluctuation in quality level under the policy. We choose the one with its simulated quality closest to the initial estimated quality as the approximation of the iteration's result.

4. Value of Information and Inventory Policy

In this section, we first establish a realistic base case. Then we simulate and compare different strategies' performances under different situations: (1) benchmark: without additional quality information and (2) improved strategies with the wholesaler owning additional quality information. We have 2 strategies in (2): (a) the wholesaler keeps it private, and the retailer estimates it by experience; (b) the wholesaler shares it with the retailer. The wholesaler's policies under (a) and (b) are solved by the MDP model we established. We compare the performance of different strategies to illustrate the value of information, analyze the optimal inventory policy under each strategy, and develop heuristic for the wholesaler.

4.1. Base Case. We take a banana supply chain in Xinfadi (one of the largest wholesale markets in Beijing) consisting of a wholesaler and a retailer as our base case example. Table 1 summarizes the problem parameters.

TABLE 1: Parameter settings.

Category	Notation	Description	Value
Global settings	α	The influence coefficient of quality	0.5
	$\bar{\beta}$	Deterioration probability for every unit at every quality level	(50%, 50%)
	A	The consumer's demand potential, a scale parameter used to determine the mean of the Poisson distribution (ton)	5
	P_0	Mean of the market price (yuan per ton)	5000
	ϕ	The probability of price keeping the same	85%
	m	The number of steps that cover the price's range	6
	γ_p	Price's step size (yuan per ton)	500
Cost parameters	c_0	Wholesaler's purchasing cost (yuan per ton)	3000
	RP	The retailer's selling price (yuan per ton)	10000
	s_W	The wholesaler's salvage value (yuan per ton)	1500
	s_R	The retailer's salvage value (yuan per ton)	500
	c_p	Penalty cost for the wholesaler (yuan per ton)	500
	K	Ordering cost (yuan)	10000
	c_h	Holding cost per unit per day (yuan per ton per day)	10

For describing the quality, the coefficient α sets the relationship between the quality's measure and the consumer's demand. In the base case, α is 0.5, which reflects a moderate degree of the consumer's sensitivity to quality as set in (2). The effect of alternative extreme values for $\alpha = 0$ and 1 will be considered in Section 5. Then we determine the deterioration probability $\bar{\beta}$ according to the characteristics of the product. The deterioration probability could be any value (no more than 100%, i.e., the model cannot be applied to the product that decays too fast). The probability in different quality levels could be different. In the base case we assume it to be (50%, 50%) to make the expected shelf life 4 days for each unit of product, which is common in the wholesale market.

In the base case one unit is one ton (1000 kg), and the unit price is also for one ton. The market price is set based on the historical data of banana market price in Xinfadi from 2013.7.1 to 2016.6.30. We set that the market price for one ton ranges from 2000 to 8000 with 5000 as mean and 500 as step size, and the probability of keeping the same price is set to be 85%. By simulation, this price setting has more than 99% price data in the range from 3000 to 7000.

In most cases, the purchasing cost 3000 is lower than the market price. On average, the wholesaler's profit margin is 40%. Based on our investigation on the retailing price, the retailer's selling price is 10000 per ton. It provides a 50% profit margin for the retailer. Since the wholesale market has a lot of potential buyers (as a secondary market) and the products are fresher than those of the retailer, the salvage value for the wholesaler's products is usually higher. We set the salvage value to be 1500 and 500 separately for the wholesaler and the retailer. In addition, the wholesaler suffers a penalty (shortage cost) of 500 for each unit of shortage.

The ordering cost happens mainly due to the long-distance transportation. From Guangzhou (south of China, near the production area of banana) to Beijing, one truck for 20,000 kg (the maximal inventory cannot exceed this volume)

requires a transportation cost of 10000, so we regard this as the ordering cost. The holding cost per unit per day is set to be 10, which includes interest, product storage, and handling costs.

4.2. Benchmark and Value of Information

4.2.1. Benchmark: No Quality Information Available. Traditionally, the products in the market are with similar quality, and the wholesaler does not take extra effort to monitor the quality. It means (1) the wholesaler makes decision only based on the total number of the inventories, (2) quality-based disposal is impossible at the wholesale stage, and (3) the wholesaler cannot control the sequence of offering products to the retailer; that is, the wholesaler's profit should be less than that with a pure FIFO sequence. Besides, the price's influence on the retailer's demand is not considered in the wholesaler's policy. We apply an (s, S) ordering policy as the real wholesaler does, focusing on the inventory state, with no disposal and with a pure FIFO sequence as a benchmark (with FIFO, it is an upper bound for this setting, and we call it the benchmark strategy). The parameters in this heuristic are obtained by searching the simulation results globally. The results are shown in Figure 2. With the quality's underlying influence considered, there could be some local optima. However, when the scale increases and it takes too much time to search globally, Figure 2 suggests us to consider a neighborhood search with larger step size to get a local optimum as an approximation to the global optimum.

4.2.2. Improved Strategies: The Wholesaler Owns Additional Quality Information. When the quality is carefully monitored, the wholesaler has additional information. The inventory policy could be improved with this information. The wholesaler can either keep this information private (denoted

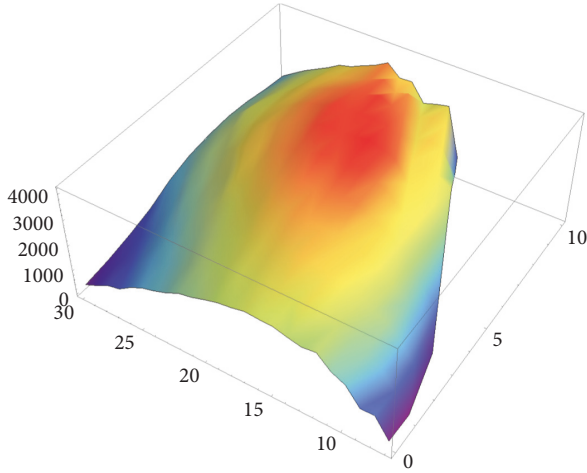
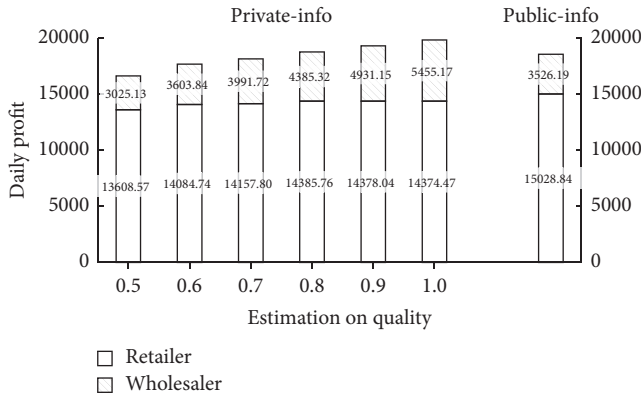
FIGURE 2: Profits over different (s, S) values.

FIGURE 3: The optimal policy's average daily profit in different strategies.

as private-info strategy) or share the information with the retailer (denoted as public-info strategy).

For the base case, the performance of the private-info strategy under the retailer's different estimations (as discussed in Section 3.3) is compared with the performance of the public-info strategy in Figure 3. In the private-info strategy, the wholesaler gets higher profit with the retailer's increasing estimation on quality. It suggests that the wholesaler builds an image of high quality (e.g., investing in the advertisement, building the brand), which encourages the retailer to order more, enlarges the wholesaler's turnover, and increases the quality in the end. However, in the long term, the retailer will learn the real quality from historical sales data. And he will adjust his ordering size to increase his profit, then the wholesaler's policy is influenced accordingly. Finally the equilibrium will fall to the one with the estimated quality close to the real quality. By simulation, the long-term quality is 0.8 for the private-info strategy, so the long-term daily profits of the wholesaler and the retailer are 4385 and 14386 separately. If the wholesaler shares the additional information

TABLE 2: Comparisons on the performance of different strategies.

Strategy	Benchmark	Private-info	Public-info
Estimated quality (constant)	0.8	0.8	—
Average quality level (θ)	0.8231	0.8248	0.8727
The wholesaler's daily profit	4176	4385	3526
The retailer's daily profit	14488	14386	15029
The retailer's daily ordering quantity	4.3862	4.3914	4.6446
The retailer's daily disposal	15.54%	16.33%	17.30%
The retailer's daily shortage	26.52%	25.89%	22.67%
The wholesaler's daily disposal	—	0.83%	2.28%
The wholesaler's daily shortage	5.43%	4.37%	3.43%
The wholesaler's daily waste	2.89%	2.25%	0.00%
The wholesaler's daily purchasing cost	12814	12999	13770
The wholesaler's daily inventory cost	88.68	88.69	68.04

with the retailer, their profits are 3526 and 15029 separately. In total, it is a little worse than the performance of the private-info strategy. What is more, the wholesaler will not choose to share the information because it leads to a big loss to his profit.

The wholesaler's optimal (heuristic) policies are compared under the three strategies more concretely by simulation, shown in Table 2.

By simulation (100 runs, each of length of 30000 periods, plus a warming-up of 300 periods), the optimal heuristic (s, S) ordering policy is $(s = 6, S = 17)$ when the quality is estimated to be 0.8. Compared with this benchmark, the wholesaler's profit increases 5.00% in the optimal private-info strategy. This increase is obtained by the wholesaler's more precise control based on the additional information: less waste, less shortage, higher quality, higher turnover, and better control with some disposal. However, the retailer's profit could decrease with the wholesaler's more precise control. In this case, the total profit of the supply chain increases with optimal policy in the private-info strategy. But for some cases with the retailer's profit margin higher than the wholesaler's, the profit of the chain could decrease comparing to the benchmark strategy when the wholesaler tries to improve his profit.

When comparing the public-info strategy with the private-info strategy, we found in public-info strategy the quality level increases by 5.81%, the retailer's profit increases,

and the shortage/waste/inventory cost decreases. The reduction of the wholesaler's profit is larger than the increase of the retailer's profit. The reason is when the retailer knows the quality information, a LIFO sequence should be applied. As a consequence, the wholesaler's profit decreases greatly with increasing disposal, while the retailer's profit only increases a little through the quality's increase.

4.3. Optimal Policy and the Heuristic. In the base case, private-info strategy performs better than the public-info strategy and the benchmark strategy. Learning from the optimal policy in private-info strategy, we come up with some heuristic rules, which could suit the reality better than the simple (s, S) policy in the benchmark strategy.

4.3.1. Overview of Optimal Policy. In Table 3, we select the most frequent states that have ordering or disposing actions under different prices to get insight into the optimal policy. We can learn the following facts from the optimal policy. (1) When the market price is low (no more than the purchasing cost), the wholesaler purchases no product from the upstream, and neither does he sell the products to the secondary market to clear his warehouse in advance. (2) The ordering quantity varies with the market price. The higher the market price is, the less the retailer orders are, and the wholesaler's ordering decreases as a result. (3) The disposal happens to the low-quality products. Mostly, it happens at the same day as ordering. And with the market price in the next day easy to forecast, it is possible for the wholesaler to dispose some products one day earlier.

The wholesaler's decision is influenced by the market price. The influence works through the retailer's decision at Level 2. In the same way, the uncertainty of the consumer's demand is absorbed by the retailer and actually has no influence on the wholesaler's decision. In this mechanism, the main uncertainty comes from the uncertainty of market price rather than the consumer's demand.

4.3.2. Heuristic Ordering Decision. As discussed, the uncertainty that the wholesaler faces, including the uncertainty of the retailer's demand, comes from the market price. Hence, for the wholesaler who makes his decision with the market price known, we could assume he makes his decision with the retailer's order known (by inferring based on the price). That is to say, a different inventory policy based on the state after sales could be applied to replace the (s, S) policy. We denote it as an (s', S') policy, where s' and S' are the reorder point and order-up-to level after sales. Also, with the reorder point and order-up-to level related to the market price, the policy is modified as $(s'(P_t), S'(P_t))$. Since $s'(P_t)$ and $S'(P_t)$ are influenced by many factors especially those related to deterioration and disposal, it is hard to find an easy approach to determine them. The following guidelines on searching could be used to obtain the reorder point and order-up-to level.

The new reorder point s' is smaller than s with the demand for the current day subtracted from s . With the retailer's ordering actually calculable in our model, it seems to be best to keep s' to be 0. However, s' could be larger than

0 because when there are some low-quality products left, we can still place a new order. We think s' is some small value; hence we regard it as a constant irrelevant to P_t .

The order-up-to level varies in different situations. After transferring the ordering quantity to order-up-to level after one day's sales, we find that when the retailer's estimation on quality is determined, all the actions under the same market price satisfy the same order-up-to level. For the retailer's different estimations on quality θ , we plot the relationship between the order-up-to level and the market price as in Figure 4.

The following features are easy to observe. (1) As a whole, with the retailer's estimation on quality decreasing, the order-up-to level decreases. (2) The relationship between order-up-to level and market price is complex. When the market price is no more than the purchasing cost, the wholesaler does not order any products and just wait for the market price to return to a normal value. (3) When the market price is larger than the purchasing cost, with the market price increasing, the order-up-to level tends to decrease, but there could be exceptions (e.g., when $\hat{\theta} = 0.8$ and $P_t = 6500$). The order-up-to level's decreasing tendency is similar to an S-curve. More precisely, its change in the low-price or high-price end is larger than that in the middle-price interval, which has a higher probability to happen with the price's mean-reverting property. It reveals that the optimal inventory policy is trying to handle the states that are highly possible to happen with a stable reaction and only makes slight adjustments at the extreme ends. It also explains why a price-irrelevant (s, S) heuristic policy could perform quite well (95.24% of the optimal): in most cases the optimal policy also gets a similar order-up-to level in the middle, while in the two ends, the order-up-to level could be different but the probability of happening is also low.

According to the guidelines above, we could develop an improved heuristic: search the reorder point irrelevant to price with low value, and search the value of order-up-to level $S'(P_t)$ with a step function; that is, search a baseline of order-up-to level for all prices higher than purchasing cost, then increase the order-up-to level at the low-price end, and do the opposite at the high-price end.

4.3.3. Heuristic Disposal Decision. The most frequent disposal states are shown in Table 4. The disposal can be classified into two categories: the disposal that happens when the ordering happens and the disposal that happens at the day without placing a new order (there are still enough products left for the next day). For those disposals with ordering, no product with low quality is left after the disposal. For the other case, there could be low-quality products left, and it depends on how many products left in total and the demand under a given market price. According to Table 4, when the wholesaler does not order, he would not dispose too many products. So we ignore this category in the heuristic for convenience.

Based on the heuristic ordering and disposal rules above, we come up with a new heuristic for the base case. With s'

TABLE 3: The most frequent states with ordering or disposing actions ($\hat{\vartheta} = 0.8$).

2000	No actions at this market price										
2500	No actions at this market price										
3000	No actions at this market price										
3500	Inv.	(3, 2)	(2, 3)	(4, 1)	(0, 0)	(1, 4)	(2, 2)	(3, 1)	(1, 3)	(5, 0)	(4, 0)
	Q	15	15	15	15	15	15	15	15	15	15
	D	0	0	0	0	0	0	0	0	0	0
	Freq.	0.42	0.38	0.24	0.20	0.17	0.12	0.11	0.06	0.06	0.04
4000	Inv.	(3, 2)	(2, 3)	(4, 1)	(1, 4)	(2, 2)	(3, 1)	(1, 3)	(5, 0)	(2, 1)	(1, 2)
	Q	15	15	15	15	15	15	15	15	15	15
	D	0	0	0	0	0	0	0	0	0	0
	Freq.	1.00	0.91	0.56	0.41	0.26	0.24	0.15	0.13	0.10	0.08
4500	Inv.	(3, 2)	(2, 3)	(4, 1)	(1, 4)	(2, 2)	(3, 1)	(2, 1)	(1, 3)	(1, 2)	(5, 0)
	Q	15	15	15	15	15	15	15	15	15	15
	D	0	0	0	0	0	0	0	0	0	0
	Freq.	1.53	1.42	0.84	0.64	0.39	0.35	0.21	0.20	0.12	0.11
5000	Inv.	(3, 2)	(2, 3)	(2, 2)	(4, 1)	(3, 1)	(1, 3)	(1, 4)	(2, 1)	(1, 2)	(4, 0)
	Q	12	12	13	12	13	13	12	13	13	13
	D	0	0	0	0	0	0	0	0	0	0
	Freq.	1.21	1.07	0.85	0.73	0.72	0.48	0.47	0.33	0.25	0.23
5500	Inv.	(2, 2)	(3, 1)	(1, 3)	(2, 1)	(4, 0)	(1, 2)	(3, 2)	(5, 4)	(2, 3)	(0, 4)
	Q	12	12	12	12	12	12	11	0	11	12
	D	0	0	0	0	0	0	0	1	0	0
	Freq.	1.60	1.24	0.97	0.39	0.34	0.32	0.26	0.23	0.23	0.21
6000	Inv.	(2, 2)	(3, 1)	(1, 3)	(2, 1)	(4, 0)	(1, 2)	(0, 4)	(7, 1)	(3, 0)	(5, 4)
	Q	12	12	12	12	12	12	12	11	12	0
	D	0	0	0	0	0	0	0	3	0	1
	Freq.	1.11	0.81	0.68	0.26	0.23	0.21	0.14	0.11	0.11	0.09
6500	Inv.	(5, 4)	(2, 2)	(3, 1)	(2, 3)	(3, 2)	(1, 3)	(4, 1)	(1, 4)	(2, 1)	(4, 0)
	Q	0	13	13	12	12	13	12	12	13	13
	D	1	0	0	0	0	0	0	0	0	0
	Freq.	0.33	0.31	0.24	0.20	0.19	0.18	0.11	0.10	0.08	0.07
7000	Inv.	(5, 4)	(2, 2)	(2, 1)	(6, 3)	(3, 1)	(1, 2)	(1, 3)	(3, 0)	(2, 3)	(4, 1)
	Q	0	11	12	0	11	12	11	12	10	11
	D	2	0	0	2	0	0	0	0	0	1
	Freq.	0.11	0.07	0.07	0.07	0.06	0.05	0.05	0.03	0.03	0.02
7500	Inv.	(4, 3)	(2, 1)	(2, 2)	(1, 2)	(1, 3)	(3, 1)	(3, 0)	(5, 4)	(0, 3)	(1, 1)
	Q	0	10	9	10	9	9	10	0	10	10
	D	1	0	0	0	0	0	0	2	0	0
	Freq.	0.03	0.02	0.02	0.02	0.01	0.01	0.01	0.00	0.00	0.00
8000	Inv.	(4, 3)	(3, 1)	(2, 1)	(2, 2)	(1, 3)	(0, 3)	(4, 0)	(0, 4)	(3, 0)	(0, 2)
	Q	0	9	10	9	9	10	10	9	10	10
	D	1	0	0	0	0	0	1	0	0	0
	Freq.	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

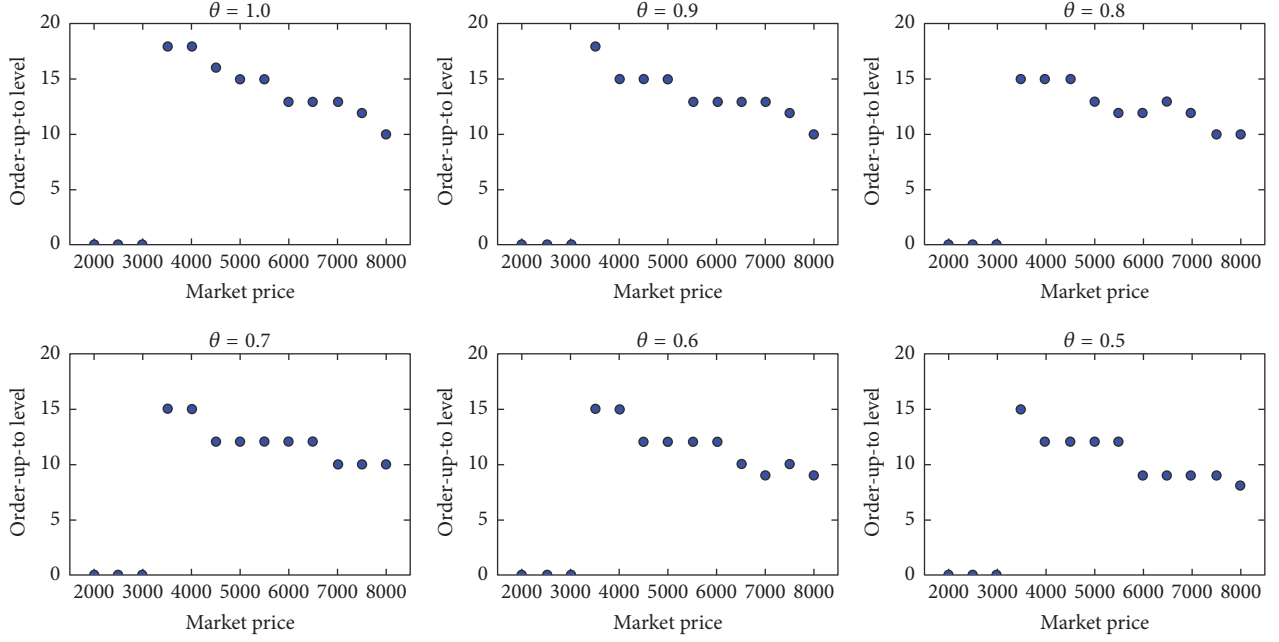


FIGURE 4: Order-up-to level versus market price.

and $S'(P_t)$ applied, we set $S'(P_t)$ to be three pieces. The two ends of $S'(P_t)$ have only 2 possible prices. Then we can solve that

$$s' = 3, \quad (11)$$

$$S' = \begin{cases} 0, & P \leq 3000, \\ 15, & 3000 < P \leq 4000, \\ 14, & 4000 < P \leq 7000, \\ 11, & P > 7000. \end{cases}$$

The wholesaler's profit is 4269, which means a 2.18% improvement comparing to the original heuristic in the benchmark strategy. With a less profit of the retailer 14331, the profit of the chain is actually worse than the original heuristic.

5. Numerical Experiments

To test the piecewise $(s', S'(P))$ inventory policy, we apply a linear design of experiments on four aspects as Table 5. For the quality related aspect, cases 1a and 1b with different α reflect the quality's different degrees of influence on consumer's demand. Cases 2a and 2b adjust the probability of deterioration. The quality level is set based on the market's response in reality, so different products could have different deterioration probability. We set deterioration probability $\bar{\beta}$ to ensure every product has an expected lifetime of 4 days (the same as that in case 0). 2a has a shorter time in high-quality stage, and 2b is on the opposite.

The market price is another important aspect. Cases 3a and 3b change ϕ , that is, the probability that price remains the same. The higher the value of ϕ is, the more stable the

TABLE 4: Most frequent disposal.

Price	Inv.	RQ	Q	D	Freq.
6500	(5, 4)	4	0	1	0.33
5500	(5, 4)	4	0	1	0.23
5500	(7, 1)	4	11	3	0.18
5000	(5, 0)	4	13	1	0.17
5000	(5, 5)	4	0	1	0.13
6000	(7, 1)	4	11	3	0.11
5000	(6, 4)	4	0	1	0.11
7000	(5, 4)	3	0	2	0.11
5000	(6, 5)	4	0	2	0.10
6000	(5, 4)	4	0	1	0.09
5000	(7, 4)	4	0	2	0.08
7000	(6, 3)	3	0	2	0.07
4500	(6, 2)	5	13	1	0.05
5000	(8, 3)	4	0	1	0.04
5500	(5, 0)	4	12	1	0.03
6500	(5, 0)	4	13	1	0.03
7500	(4, 3)	3	0	1	0.03
7000	(4, 1)	3	11	1	0.02
6500	(8, 1)	4	12	4	0.02
5000	(5, 1)	4	12	1	0.02

price is. Cases 4a and 4b make the price's range unchanged, but change the price's variance: in case 4a, with larger step size but less steps, the variance is increased. In case 4b, the variance decreases by an opposite adjustment.

The following experiments 5a, 5b, 6a, and 6b are designed to show the influence of profit margin: in cases 5a and 5b, the wholesaler's profit margin is changed by changing the

TABLE 5: Design of experiments.

	α	$\vec{\beta}$	ϕ	m	γ_p	c_0	RP	s_w	s_r
Base case									
0	0.5	(50%, 50%)	0.85	6	500	3000	10000	1500	500
Quality									
1a	1								
1b	0								
2a		(36%, 83%)							
2b		(83%, 36%)							
Market price									
3a			0.95						
3b			0.75						
4a				4	750				
4b				8	375				
Profit margin									
5a						4000			
5b						2000			
6a							12000		
6b							8000		
Salvage value									
7a								2000	
7b								1000	
8a									1000
8b									0

purchasing cost (on average, 20% and 60% profit margin for 5a and 5b). In cases 6a and 6b, the retailer's profit margin varies with different selling price of the retailer (60% and 40% on average for 6a and 6b).

The last aspect is on salvage value: cases 7a and 7b change the salvage value to the wholesaler, and cases 8a and 8b change the salvage value to the retailer.

Later, we apply the basic (s, S) policy and the modified piecewise $(s', S'(P))$ to each case. Then we compare them with the optimal outcome from our bilevel model to show the performance of the new heuristic in different scenarios.

The performance of different heuristics compared to the performance of optimal policy is shown in Table 6. In the public-info strategy, the wholesaler's profit is always less than that in a private-info policy, so those are not shown in Table 6. The original (s, S) policy performs well. On average, the wholesaler gets a 94.70% performance of the optimal policy in the private-info strategy (except that in case 5a all heuristic policies result in a negative profit), while the piecewise $(s', S'(P))$ policy has an improvement of 2.40% on average compared to (s, S) . We get (s, S) and $(s', S'(P))$ by global search, and the output shows no obvious rules on their values. But as discussed, we can obtain a local optimum as the result by neighborhood search.

Cases 1a and 1b change the degree of the quality's influence on the consumer's demand. Both cases fall into the quality level of 0.8, but actually in case 1a the quality is slightly lower than that in case 1b (0.82 versus 0.83 under the optimal policy), even the customers are more sensitive in case 1a. With consumers being more sensitive, the retailer

acts more conservatively. Then the wholesaler's turnover decreases, which decreases the quality. In case 2a, the (s, S) policy may fall to the quality level of 0.7, which is worse than the 0.8 in the optimal policy. It means for the products with short high-quality stage (high-quality product deteriorates faster), a precise control could improve the quality perception efficiently. In case 2b, when the low-quality stage is short, the qualities in the heuristic and the optimal policy are all 0.9. In cases 2a and 2b, the optimal policy improves the wholesaler's profit greatly. The reason is that when one quality stage is short, it costs less to control the quality by ordering and disposing.

Cases 3a and 3b reflect that the wholesaler obtains higher profit with more stable price (more likely to remain the same). The influence is not significant. Cases 4a and 4b represent that higher variance leads to lower profit. For the high-variance scenario, the piecewise heuristic gets a significant improvement from the original one. In case 4a with price's intensive fluctuation, the modified heuristic performs quite well and it leads to a profit higher than the optimal policy due to the error of simulation. In case 4b, we also observe that the order-up-to level of different price has a larger gap with the larger variance.

Cases 5a, 5b, 6a, and 6b show that profit margin of both the wholesaler and retailer has a great influence on the profit. Case 5a shows that, for low profit margin scenario, when the original (s, S) policy is not profitable, the precise control on inventory could be profitable. In cases 6a and 6b, the wholesaler's profit is influenced by the retailer's profit margin through the wholesaler's turnover. When the retailer's profit

TABLE 6: Comparisons of different policies.

	Optimal policy (private-info)		(s, S) heuristic (% of optimal policy)			(s', S'(P)) heuristic (% of optimal policy)		
	$\hat{\theta}$	Profit	$\hat{\theta}$	Policy	% of optimal policy	$\hat{\theta}$	Policy	% of optimal policy
0	0.8	4385.32	0.8	(6, 17)	95.24	0.8	(3, 15-14-11)	97.34
1a	0.8	3602.92	0.8	(6, 16)	93.84	0.8	(2, 16-12-9)	98.92
1b	0.8	5490.44	0.8	(7, 20)	96.68	0.8	(3, 18-15-12)	99.00
2a	0.8	3277.53	0.7	(6, 17)	92.01	0.7	(1, 13-13-10)	93.61
2b	0.9	5508.42	0.9	(7, 20)	94.35	0.9	(2, 17-16-13)	94.66
3a	0.8	4445.22	0.8	(7, 18)	93.86	0.8	(1, 15-13-9)	95.20
3b	0.8	4385.12	0.8	(6, 18)	94.80	0.8	(2, 16-14-12)	96.91
4a	0.8	4203.88	0.8	(7, 18)	93.50	0.8	(2, 15-13-12)	100.94
4b	0.8	4498.65	0.8	(7, 18)	95.06	0.8	(2, 18-15-11)	95.19
5a	0.8	457.28	No profitable policy			0.8	(1, 13-12-11)	94.13
5b	0.8	9015.69	0.8	(7, 21)	96.43	0.8	(3, 21-16-12)	97.37
6a	0.8	5400.04	0.8	(8, 20)	97.18	0.8	(3, 16-15-14)	99.18
6b	0.8	3035.84	0.8	(5, 16)	92.49	0.8	(1, 15-12-8)	98.64
7a	0.8	4437.72	0.8	(6, 17)	94.11	0.8	(3, 15-14-11)	96.48
7b	0.8	4401.77	0.8	(6, 17)	94.88	0.8	(3, 15-14-11)	96.69
8a	0.8	4849.45	0.8	(7, 20)	95.41	0.8	(2, 16-15-11)	97.05
8b	0.8	4399.70	0.8	(6, 17)	95.36	0.8	(3, 17-14-12)	96.37

margin is low, the wholesaler's profit could be significantly improved by applying the modified policy.

Cases 7a, 7b, 8a, and 8b show that the salvage value has slight influence on the wholesaler's profit, which means the disposal is not the crucial decision with a FIFO offering sequence. The wholesaler's salvage value has little influence on the wholesaler's profit, but for the retailer, when increasing his salvage value, the profit may be promoted through the increase of the retailer's ordering size.

The observations lead to three main managerial insights.

(1) In a wholesaler-retailer system, keeping the wholesaler's additional quality information private could decrease the influence of quality's fluctuating on the retailer's ordering, which makes the demand more stable and increases the profit of the chain. (2) To include the influence of price's fluctuating, the wholesaler's decision changes; hence his profit increases. The greater the fluctuating is, the more efficient the improved strategy should be. (3) The influence of salvage value is so significant as the influence of the profit margin. Based on these insights, the new inventory policy under private-info strategy could include the influence of market price and improve the wholesaler's profit in all cases significantly.

6. Concluding Remarks

In this paper, we introduce the stochastic market price's influence into a wholesaler-retailer two-level system in the wholesale market. In this system, the wholesaler directly faces the retailer's demand, which is influenced by the stochastic market price, an exogenous random variable with mean-reverting property. The retailer's demand is also influenced

by the consumer's stochastic demand, which is influenced by the quality. We establish a bilevel stochastic dynamic programming model on this complex system to help the wholesaler making better decisions. We also notice that the quality's difference that influences the consumer's demand may not be significantly shown at the wholesale stage. By new technology monitoring the relevant parameters, the wholesaler gets additional information on the quality. We answer how much value this information has and what strategy the wholesaler should apply to make use of it. We also analyze the structure of the optimal policy and develop a piecewise inventory policy with the consideration on the additional quality information and the influence of stochastic market price.

We have the following findings:

- (i) In the market, the quality seems to be the same, but it is still important for the wholesaler to monitor the quality more precisely. The wholesaler can make better inventory decisions with additional quality information.
- (ii) However, it is better not to share this information with the retailer, which will significantly reduce the wholesaler's profit and usually results in a worse performance for the chain.
- (iii) In the wholesale market where the price is the main source of uncertainty, it is helpful to include the consideration on price into the inventory policy.
- (iv) The wholesaler could dispose low-quality product to the secondary market with additional information. In the optimal policy, the disposal happens either

at the same day as the ordering or one day earlier than ordering. This behavior mode relies on how the retailer makes decision, with which the market price is the main uncertainty. For convenience, we could only consider the disposal happening at the same day as ordering.

- (v) With the demand of the current day directly estimated from the market price, we could regard the wholesaler actually makes decision after the sales. Then the optimal inventory policy is similar to a (s, S) policy, with the reorder point and order-up-to level related to price. Since the reorder point becomes low in the new decision time, we fix it to a constant. The order-up-to level is related to price and shows an S-shape. Hence we approximate it with a piecewise function and get an $(s', S'(P))$ inventory policy, which has a 2.40% improvement on average comparing to the original (s, S) .
- (vi) In the bilevel system, the quality's influence is complicated. The wholesaler influences the quality directly. The retailer estimates the quality from the historical sales data and is less sensitive to the quality change. But the retailer's decision is usually conservative. As a result, the system could also be more sensitive to the small disturbance which happens to the retailer.

The problem is still not fully studied. First, with the MDP model hard to solve for the multichain structure problem, we only apply a two-level quality system and solve the problem with a quantity no more than 20. It reflects some features but limits the system's change to be less smooth. Second, we describe the price's feature based on an Ehrenfest chain, which indeed has influences on the wholesaler's decision. But due to the limitation of the scale, we only set a few possible values of the price, which limits our discussion on the influence of the market price's feature. For further exploration, an approximate dynamic programming model could be promising when facing to a larger scale problem with more specific features included.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] A. J. Watson, "Conflict over Cabbages: the reform of wholesale marketing in China," in *In The Third Revolution in the Chinese Countryside*, pp. 144–163, Cambridge University Press, 1996.
- [2] B. Li, "Price reform for agricultural products," in *In The Third Revolution in the Chinese Countryside*, pp. 81–88, Cambridge University Press, 1996.
- [3] R. Garnaut, Guo Shutian, and Ma Guonan, "The third revolution in the Chinese countryside," *The third revolution in the Chinese countryside*, 1996.
- [4] B. A. Kalyon, "Stochastic prices in a single-item inventory purchasing model," *Operations Research*, vol. 19, pp. 1434–1458, 1971.
- [5] P. Berling and V. Martínez-de-Albéniz, "Optimal inventory policies when purchase price and demand are stochastic," *Operations Research*, vol. 59, no. 1, pp. 109–124, 2011.
- [6] Y. Chen, W. Xue, and J. Yang, "Technical note - Optimal inventory policy in the presence of a long-term supplier and a spot market," *Operations Research*, vol. 61, no. 1, pp. 88–97, 2013.
- [7] K. G. Grunert, "Food quality and safety: consumer perception and demand," *European Review of Agricultural Economics*, vol. 32, no. 3, pp. 369–391, 2005.
- [8] M. Bertolini, G. Ferretti, G. Vignali, and A. Volpi, "Reducing out of stock, shrinkage and overstock through RFID in the fresh food supply chain: Evidence from an Italian retail pilot," *International Journal of RF Technologies: Research and Applications*, vol. 4, no. 2, pp. 107–125, 2013.
- [9] A. Kirmani and A. R. Rao, "No pain, no gain: A critical review of the literature on signaling unobservable product quality," *Journal of Marketing*, vol. 64, no. 2, pp. 66–79, 2000.
- [10] J. E. Hobbs, "Information asymmetry and the role of traceability systems," *Agribusiness*, vol. 20, no. 4, pp. 397–415, 2004.
- [11] M. Ferguson and M. E. Ketzenberg, "Information sharing to improve retail product freshness of perishables," *Production Engineering Research and Development*, vol. 15, no. 1, pp. 57–73, 2006.
- [12] R. Haijema, "Optimal ordering, issuance and disposal policies for inventory management of perishable products," *International Journal of Production Economics*, vol. 157, no. 1, pp. 158–169, 2014.
- [13] B. Colson, P. Marcotte, and G. Savard, "Bilevel programming: a survey," *4OR*, vol. 3, no. 2, pp. 87–107, 2005.
- [14] D. van Dijk, E. M. T. Hendrix, R. Haijema, R. A. Groeneveld, and E. C. Van Ierland, "On solving a bi-level stochastic dynamic programming model for analyzing fisheries policies: Fishermen behavior and optimal fish quota," *Ecological Modelling*, vol. 272, pp. 68–75, 2014.
- [15] M. L. Puterman, *Markov decision processes: discrete stochastic dynamic programming*, John Wiley & Sons, 2014.

