Selling Surplus Foods Through Surprise Bags: Online Inventory Routing Across Dual Channels

Yiding Feng
Hong Kong University of Science and Technology, Hong Kong SAR, China. ydfeng@ust.hk

Chui Shan Lee
Hong Kong University of Science and Technology, Hong Kong SAR, China. tracylcs@ust.hk

Guodong Lyu
Hong Kong University of Science and Technology, Hong Kong SAR, China. imlyu@ust.hk

Huan Zheng
Shanghai Jiao Tong University, Shanghai, China. zhenghuan@sjtu.edu.cn

Abstract.

Reducing food waste while preserving profit is an important operational problem for retailers of perishable products. We study a finite-horizon inventory allocation problem with a regular channel and an opaque surprise-bag channel. Before observing each period's regular demand, the retailer chooses product display quantities and surprise-bag offers. Each offered bag sells and creates a noncancellable commitment that must be fulfilled from inventory remaining at the horizon. The model incorporates waste costs and linear constraints on the expected aggregate composition of bag allocations, while allowing arbitrary correlation in regular demand across products and periods. Under productwise regular-channel profitability, an LP-based reservation policy achieves the optimal competitive ratio $1/(2+\kappa)$ against a clairvoyant benchmark, where $\kappa = w/s$ is the waste-cost-to-bag-reward ratio. An adaptive reservation policy weakly dominates the static policy in expected profit on every instance and inherits its guarantee. We also identify sufficient conditions under which deferred product assignment strictly improves optimal expected profit over immediate assignment, and extend the static guarantee to fixed multi-unit bags. In simulations calibrated from cafeteria sales data, the adaptive policy improves mean profit over the static policy by 6.2%–9.2% across the tested quality-floor settings.

1. Introduction

Reducing food loss and waste is central to the 2030 Agenda for Sustainable Development, yet progress remains far off track despite sustained efforts across the global food supply chain. A substantial share of this waste originates not only in fields and warehouses but also at the final link: retail and food-service operations. Restaurants, bakeries, grocers, and other operators must stock or prepare perishable products before customer demand is fully known [Akkaş and Gaur, 2022; Nahmias, 1982]. This timing mismatch becomes especially costly near the end of the selling horizon, when products that were valuable earlier now have only a short window left to be sold. The resulting surplus is substantial. ReFED [2024] reports that U.S. retailers generated 3.98 million tons of surplus food in 2024, with a quarter sent to landfill or incineration.

Surplus-food platforms such as *Too Good To Go* have emerged as one response to this problem [Too Good To Go, 2024]. Through these platforms, retailers can sell discounted "surprise bags" composed of surplus food that might otherwise be discarded [Yang and Yu, 2025; Zhou et al., 2024]. For consumers, the

channel offers a low-priced way to purchase food, at the cost of reserving a bag without knowing its exact contents until pickup. For retailers, the same opacity is an operational lever: because the contents are not specified when the bag is sold, the bag can later be assembled from whichever products remain available [Yang and Yu, 2025; Zhou et al., 2024].

The operational problem, however, is more subtle than simply packing whatever is left at closing. Bags must be made available while the day is still unfolding, before the retailer knows exactly which products will remain. Each bag commitment therefore creates an immediate fulfillment obligation, but product assignment can be deferred: the retailer sells an opaque bag now and later determines which physical inventory will satisfy it. Reserving more bags can secure revenue and reduce the risk of waste, but it also leaves less inventory available for customers who may arrive later. Thus, surprise-bag platforms create not only an end-of-day surplus-disposal problem, but an online inventory-routing problem across both regular demand and surprise-bag demand channels. This raises the central question we address:

Is there an inventory-value measure that leads to a simple reservation framework with a best-possible expected-profit guarantee, while still accommodating a broad class of operational constraints?

Surprise-bag marketplaces involve several intertwined platform-design choices, and recent work has begun to study some of these dimensions. Yang and Yu [2025] analyze surprise clearance as a selling scheme and compare it with no clearance and transparent clearance. Zhou et al. [2024] study surprise-bag design, focusing on pricing, bag-value allocation, supplementation, and reputation dynamics. Dong et al. [2025] study platform-level market design and examine how pricing policies affect seller participation and waste diversion. Together, these papers motivate surprise-bag marketplaces as an important operational setting. We isolate an understudied store-level inventory-routing problem created by surprise bags, in which bag timing and fulfillment must be coordinated with regular sales as demand is gradually revealed. Since the selling scheme, pricing, and other marketplace-design are typically set before the selling day or managed outside the retailer's intraday control problem, we treat them as primitives rather than additional controls.

1.1. Our Contributions

With this scope in place, we formulate a finite-horizon model of online inventory routing across a regular channel and a surprise-bag channel. The selling day is divided into discrete periods, and the retailer starts with initial inventory of multiple perishable products. Demand in the regular channel is stochastic with a known distribution, but its realization is revealed only over time. For the surprise-bag channel, we assume demand is abundant relative to the retailer's bag supply, but a per-period offer capacity limits the total number of surprise bags that can be offered.

In each period, the retailer chooses regular display quantities for each product and the number of surprise bags to offer. Both decisions are subject to operational limits, such as display capacity in the regular channel and offer capacity in the surprise-bag channel. Displayed products are sold whenever regular demand materializes, while each offered bag becomes a noncancellable fulfillment commitment. This creates the routing

tradeoff: a unit sold through the regular channel cannot be used to fulfill a surprise bag, while each bag commitment reduces the inventory available for later regular sales. The retailer need not specify the product inside the bag when it is offered, but each bag must eventually be fulfilled by one unit of physical inventory.[1] Thus, each unit ultimately has one of three outcomes: it is sold through the regular channel, assigned to a surprise bag, or discarded as leftover inventory. The objective is to maximize expected profit, accounting for regular sales revenue, surprise-bag revenue, and the waste-handling cost of leftover inventory.

Our formulation also allows retailers to control the aggregate product mix allocated to surprise bags. We represent these requirements by linear constraints on expected product allocations, with expectations taken over demand and policy randomization. The constraints can impose an average-quality floor, upper bounds on the expected allocation of individual products, and limits on category concentration [e.g., El Housni et al., 2024; Fay et al., 2015; Gao et al., 2025; Khan and DePaoli, 2024; Rosales et al., 2019; Sun et al., 2024, 2025; Tian et al., 2024; Zhou et al., 2024]. Section EC.2 gives their explicit forms.

From a theoretical perspective, our model is closely related to the prophet inequality [Samuel-Cahn, 1984], the classic framework for studying Bayesian online resource allocation. We highlight several key differences between our model and the standard prophet inequality setting. First, in our model the rewards for both regular sales and surprise bags are fixed over time, while the joint demand distribution allows arbitrary correlation across products and periods; the prophet inequality typically assumes a single unit demand per round with rewards drawn from independent distributions. Second, the decision variables differ: in our model the decision maker chooses per-period supply quantities for the regular and surprise-bag channels, subject to remaining inventory and a final bag-composition constraint that must be satisfied at the end of the horizon, whereas in the prophet inequality the decision maker chooses whether to allocate each arriving demand unit. Third, while most prophet inequality work focuses on reward maximization (with only a few exceptions Ekbatani et al., 2024; Livanos and Mehta, 2024 studying cost minimization or mixed-sign profit maximization), our model incorporates both revenue and waste costs, yielding a mixed-sign objective together with general bag composition constraints. This mixed-sign objective is tied directly to the food-waste application: leftover inventory is not merely unsold revenue, but also carries a waste-handling cost. This is widely regarded as a more challenging paradigm for online algorithm design [Babaioff et al., 2009; Ekbatani et al., 2023, 2024]. Finally, our algorithmic framework differs from the classic prophet inequality approach; we discuss our proposed policies below.

LP-based reservation policy. Our first algorithmic result shows that a simple beginning-of-day reservation policy achieves the best possible worst-case guarantee among online policies. The policy commits at the

[1] Some surprise-bag offerings consist of one physical item whose variety is unspecified at purchase. For example, Too Good To Go listings for Pizza My Heart in Sunnyvale, California, and Polly's Pies Restaurant & Bakery in Montebello, California, offer a whole pizza with unspecified variety or flavor and one whole pie from a selection of varieties, respectively. A unit may also represent an exogenously specified, preassembled bundle, provided it is treated as an indivisible inventory unit in both channels. Our one-unit base model also extends to surprise bags containing exactly m physical units. In Section EC.4, we study this extension model and generalize our algorithmic results.

start of the horizon to a fixed reservation split, reserving some inventory for future surprise-bag fulfillment while leaving the rest available for regular demand. This split must balance the bag reward from reserved units against the opportunity cost of removing those units from regular sales.

Our algorithm, the *LP-based reservation policy* (Algorithm 2), contains two main technical ingredients: (i) A linear program $\mathrm{LP}_{\mathrm{st}}$ that yields a fractional canonical solution splitting inventory between the regular channel and the surprise-bag channel. (ii) A *lossless* dependent-rounding procedure produces an integral reservation before the selling horizon. Its suffix structure permits online implementation without loss in expected profit, while preserving aggregate composition constraints in expectation and respecting bag capacity on every realization. A crucial concept in the formulation of the $\mathrm{LP}_{\mathrm{st}}$ is the *regular value*, which quantifies the opportunity cost of removing a unit from regular sales. More specifically, it is defined as the expected net contribution of leaving the unit available to regular demand. In the LP-based reservation policy, $\mathrm{LP}_{\mathrm{st}}$ uses these regular values to rank the reservation tradeoffs across units, selecting bag reservations only when the bag reward and composition benefits justify the lost regular-sales contribution. We adapt the dependent rounding scheme [Chekuri et al., 2010; Gandhi et al., 2006] to a suffix-form optimal reservation plan. The suffix structure preserves the unit-level value accounting after rounding, while marginal preservation ensures the expected composition requirements.

Despite accommodating general soft linear constraints on expected bag composition, under the mild productwise regular-channel profitability (Assumption 1) assumption, the LP-based reservation policy achieves the tight $1/(2+\kappa)$ competitive ratio against a benchmark that knows future demand, where κ is the ratio of waste cost to bag reward. To obtain this guarantee, we first introduce a uniform-gain LP, $\mathrm{LP}_{\mathrm{ug}}$, which provides an upper bound on the profit of the optimal offline benchmark. We then use a dual-fitting argument to bound the ratio between $\mathrm{LP}_{\mathrm{st}}$—whose value equals the expected profit of our policy thanks to the lossless rounding—and $\mathrm{LP}_{\mathrm{ug}}$. Intuitively, the dual-fitting argument shows that the offline policy's hindsight advantage can be charged to the same regular value tradeoff used by $\mathrm{LP}_{\mathrm{st}}$. To complement this positive result, we prove a matching lower bound: no online policy can improve this ratio in general, even in a one-product, one-period instance. This shows that the loss is intrinsic to committing before regular demand is revealed.

At first glance, our proposed LP-based reservation policy shares similarities with many online algorithms in the prophet inequality literature [e.g., Alaei et al., 2012; Feldman et al., 2016; Feng et al., 2024], which first solve a linear program (e.g., an ex ante relaxation or fluid approximation) and then conduct online rounding (e.g., an online contention resolution scheme). However, our approach differs in a crucial respect. As noted above, $\mathrm{LP}_{\mathrm{st}}$ is *not* an upper bound on the optimal offline benchmark. Instead, our dependent rounding procedure provides a lossless online implementation of its solution. Because our rounding is lossless, our analysis does not need to bound rounding loss, which is the standard step when solving an ex ante relaxation and then applying an online contention resolution scheme. Rather, we use a dual-fitting argument to

bound the gap between $\mathrm{LP}_{\mathrm{st}}$ and the uniform-gain LP $\mathrm{LP}_{\mathrm{ug}}$, which relaxes the optimal offline benchmark. Thus, the LP-based structure of the policy differs from the usual relaxation-and-rounding approach.

Adaptive reservation through re-solving. Our second algorithmic result develops an adaptive reservation policy (Algorithm 3) that weakly dominates the LP-based reservation policy in expected profit on every instance. At the beginning of each period, the policy re-solves the residual linear program $\mathrm{LP}_t(S^{(t-1)})$ (analogous to $\mathrm{LP}_{\mathrm{st}}$) on the remaining inventory, using updated residual regular values and residual soft budgets, as well as the conditional distribution of remaining demand. The same regular-value principle reappears in residual form: at each re-solving step, the residual LP updates the residual regular values from the current residual inventory and remaining horizon, then uses them to re-rank the reservation tradeoffs under the residual bag capacity and soft budgets. The resulting fractional reservation plan is rounded into a remaining-horizon schedule, whose current-period component determines the bag offers and regular displays for that period. After demand is realized, all uncommitted surviving units are pooled again, the conditional distribution of remaining demand is updated, and the policy re-solves for the next period.

The main technical challenge is preserving feasibility while repeatedly updating the future schedule after observing demand. Although the dominance relation is non-strict, even this guarantee is not immediate in our model. Each residual-LP solution must be rounded into an integral remaining-horizon schedule while preserving the general expectation-level bag-composition constraints. The rounding must also be lossless for the residual-LP value: its current component determines feasible current-period bag offers and regular displays, its unexecuted tail certifies feasibility of the next residual LP, and its unit marginals support the telescoping value comparison. We address this by carrying forward residual soft budgets and ensuring that the rounded schedule's unexecuted tail remains feasible for the next residual LP. This tail-feasibility property (Lemma 5) yields a one-step residual-LP value comparison, which telescopes over time. This gives instance-wise expected-profit dominance over the LP-based reservation policy; consequently, the adaptive reservation policy inherits the same $1/(2+\kappa)$ competitive guarantee.

Value of deferred product assignment. Our model permits a retailer to accept a surprise-bag order before deciding which product will fulfill it. We distinguish this deferred assignment (DA) from immediate assignment (IA), under which each accepted bag is irrevocably assigned a product unit before current-period regular demand is observed. Although DA can replicate any IA policy, its additional flexibility need not increase optimal expected profit: waiting to accept a bag can eliminate the need for reassignment, and reassignment may not change the best attainable profit. In Section 5, we identify sufficient conditions for a strict DA advantage in symmetric instances with independent uniform demands and no composition constraints. The conditions combine limited per-period bag-offer capacity with an intermediate range of demand. Exact dynamic programming calculations under uniform and binomial demand illustrate how the DA advantage varies with demand and bag capacity. This comparison is distinct from the value of adaptive reservation:

both proposed reservation policies protect product-specific inventory for accepted orders, and their performance guarantees do not require reassignment of those orders.

Numerical experiments calibrated from sales data. We evaluate the reservation policies in simulations calibrated from transaction-level sales data from a corporate cafeteria operated by a Chinese catering company. Across the tested average-quality requirements, both policies outperform the uniform-share and certainty-equivalent heuristics in mean profit, and adaptive reservation improves on static reservation by 6.2%–9.2%. Without a quality floor, the adaptive policy achieves 88.4% of the simulated offline benchmark's mean profit. With positive quality floors, it achieves 88.2%–89.3% of an ex ante upper bound. These results illustrate the benefit of revising reservations as inventory evolves. They are obtained under a demand model calibrated from observed sales, rather than from a field implementation of the policies. Section 6 details the calibration, benchmarks, and evaluation protocol.

For a more in-depth discussion of related work, see Section EC.1.

2. Preliminaries

We consider a retailer managing a fixed inventory of perishable products over a finite selling horizon. The inventory consists of $N \in \mathbb{N}_{>0}$ product types, with $q_i \in \mathbb{N}_{>0}$ units of product $i \in [N]$ available initially. By the end of the selling horizon, each unit is either sold through the regular channel, used in a surprise bag sold through the surprise-bag channel, or discarded. A regular sale of product i earns $r_i \in \mathbb{R}_{>0}$; each surprise bag sold earns $s \in \mathbb{R}_{>0}$, regardless of the product allocated; and each discarded unit incurs a waste cost[2] $w \in \mathbb{R}_{\geq 0}$. Motivated by practical applications, we assume $r_i > s \geq w$ for every product $i \in [N]$.

The selling horizon is divided into $T \in \mathbb{N}_{>0}$ periods, indexed by $t \in [T]$. In each period t, let $d_{i,t} \in \mathbb{N}$ denote the demand for each product $i \in [N]$ in the regular channel. The full demand sequence $\{d_{i,t}\}_{i\in[N],t\in[T]}$ is drawn from a joint distribution $\mathcal{D}$, which may exhibit arbitrary correlation across products and periods. At the start of the horizon, the retailer knows the joint distribution $\mathcal{D}$, but not the realized demands $\{d_{i,t}\}_{i\in[N],t\in[T]}$. At the beginning of each period t, let $q_{i,t-1} \in \mathbb{N}$ denote the remaining physical inventory of product i, with $q_{i,0} \triangleq q_i$. Before observing period-t demand, the retailer makes two decisions: how many units of each product to display in the regular channel and how many surprise bags to make available.

First, for each product $i \in [N]$, the retailer chooses a display quantity $x_{i,t} \in \mathbb{N}$ for the regular channel. The display quantity cannot exceed the remaining physical inventory of the product, and is also subject to an exogenous regular-display cap $X_i \in \mathbb{N}$. Thus, $x_{i,t} \leq \min\{q_{i,t-1}, X_i\}$. The display cap captures per-period limits on making a product available for regular sales, such as display space or handling capacity. Units not displayed remain in inventory and are not exposed to regular demand in that period.

[2] We use a product-independent waste cost to capture common food-waste handling and disposal costs.

Second, the retailer chooses the number of surprise bags $y_t \in \mathbb{N}$ to make available in period t. To maintain a scarce offering, the retailer makes only a limited number of bags available, all of which are assumed to sell. An exogenous bag-offer cap $Y \in \mathbb{N}$ bounds the number of bags that can be made available in each period. The decision must satisfy[3] $y_t \le Y$.

After the retailer makes both decisions, period-t regular demand is realized. The realized demand profile is then observed. Regular sales of product i are limited by the quantity displayed and therefore equal $\min\{d_{i,t}, x_{i,t}\}$. Each bag sold must be fulfilled with one unit at the end of the horizon; the product used to fulfill it is selected only then. Thus, the physical inventory of product i remaining after period t is $q_{i,t} = q_{i,t-1} - \min\{d_{i,t}, x_{i,t}\}$.

After period T, the retailer assigns the remaining inventory to the surprise bags sold over the horizon. For each product $i \in [N]$, let $z_i \in \{0, \dots, q_{i,T}\}$ denote the number of units used for this purpose. Each bag sold requires exactly one allocated unit. Hence, the total number of allocated units must satisfy

$$\sum\nolimits_{i\in[N]} z_i = \sum\nolimits_{t\in[T]} y_t.$$

The product mix used to fulfill sold surprise bags may also be subject to composition requirements. We represent these requirements by $L \in \mathbb{N}$ linear constraints on the expected product allocation across sold surprise bags. Unless otherwise specified, all expectations in this paper are taken over regular-demand realizations and any randomization used by the policy under consideration. For each row $\ell \in [L]$, let $b_\ell \ge 0$ denote its soft budget and, for each product $i \in [N]$, let $a_{i,\ell} \in \mathbb{R}$ specify the contribution of one unit of product i to its usage. A feasible policy must satisfy

$$\sum\nolimits_{i\in[N]} a_{i,\ell}\mathbb{E}[z_i] \le b_\ell, \qquad \ell \in [L]. \tag{1}$$

We call b_ℓ the soft budget of row ℓ and the left-hand side its expected soft-budget usage. We call these constraints soft because they are imposed on the expected allocation across surprise bags sold rather than on every realized allocation. We require $b_\ell \ge 0$, which ensures that the zero allocation, and therefore making no surprise bags available, satisfies every composition requirement.[4] This formulation captures several useful composition requirements; Section EC.2 gives their explicit linear forms:

- **Average-quality floor.** A lower bound on the average product quality of the expected allocation, measured by a product-specific quality score [Sun et al., 2024, 2025; Zhou et al., 2024];
- **Product-availability caps.** Upper bounds on the expected number of units of individual products assigned to sold surprise bags [Gao et al., 2025; Khan and DePaoli, 2024; Rosales et al., 2019];

[3] The model and analysis can be extended to time-dependent caps by replacing X_i and Y with $X_{i,t}$ and Y_t, respectively.

[4] Allowing $b_\ell < 0$ leads to fundamental impossibility: even a single such constraint can render no positive performance guarantee achievable; see Section EC.3.

- **Category-concentration limits.** Upper bounds on the fraction of the expected allocation that may come from a designated subset of products [El Housni et al., 2024; Fay et al., 2015; Tian et al., 2024].

Together, the inventory, economic parameters, demand distributions, operating limits, and composition requirements defined above specify an instance

$$I \triangleq \Big(N, T, L,\ \{q_i, r_i, X_i\}_{i\in[N]},\ s, w, Y,\ \mathcal{D},\ \{\{a_{i,\ell}\}_{i\in[N]},\ b_\ell\}_{\ell\in[L]}\Big).$$

Given an instance I, an online policy may randomize independently of the entire demand process, but its period-t display and bag-availability decisions may depend only on the known instance, the observed demand and action history through period $t-1$, and a period-t random seed independent of the demand process and the policy's preceding random seeds. Its end-of-horizon allocation may use all information revealed over the horizon. We call a policy feasible if, surely, its display and bag-availability decisions satisfy the inventory and cap constraints and its sold bags can be fulfilled, while the composition requirements hold in expectation.

Over the horizon, product i has total regular sales $q_i - q_{i,T}$, while $q_{i,T} - z_i$ units remain unused and are discarded at the end of the horizon. For a feasible policy Π on instance I, define its realized profit as

$$\textsc{Profit}[\Pi; I] \triangleq \sum\nolimits_{i\in[N]} r_i\big(q_i - q_{i,T}\big) + s\sum\nolimits_{t\in[T]} y_t - w\sum\nolimits_{i\in[N]}\big(q_{i,T} - z_i\big).$$

The retailer seeks to maximize $\mathbb{E}[\textsc{Profit}[\Pi; I]]$.

We evaluate a feasible online policy Π relative to an offline optimal policy. This policy satisfies the same feasibility requirements, but observes the full demand realization $\{d_{i,t}\}_{i\in[N],\,t\in[T]}$ before making any decisions and maximizes expected profit among all such policies. Let OPT denote this policy; its expected profit is our offline benchmark.

Consider the regular-only greedy policy, under which the surprise-bag channel is unavailable. Let RO denote this policy. In every period, the retailer displays as many units of each product as the remaining physical inventory and the regular-display cap allow. Let $\chi_i \in \{0, \ldots, q_i\}$ denote the total regular sales of product i under this policy. If inventory remains available, product i can serve up to $\min\{d_{i,t}, X_i\}$ units of demand in period t. Hence, for any demand realization,

$$\chi_i = \min\left\{q_i,\ \sum\nolimits_{t\in[T]} \min\{d_{i,t}, X_i\}\right\}.$$

The expected profit contributed by product i is then

$$\mathbb{E}[\textsc{Profit}_i[\text{RO}; I]] \triangleq r_i\mathbb{E}[\chi_i] - w\big(q_i - \mathbb{E}[\chi_i]\big). \tag{2}$$

ASSUMPTION 1 **(Productwise regular-channel profitability).** $\mathbb{E}[\textsc{Profit}_i[\text{RO}; I]] \geq 0$ *for all* $i \in [N]$.

Economically, Assumption 1 restricts attention to instances in which every product is non-loss-making through the regular channel alone. A violation may reflect upstream overproduction relative to regular demand and display capacity. We leave such instances outside the scope of our analysis.

DEFINITION 1 (α-COMPETITIVE POLICY). A feasible online policy Π is α-competitive, for $\alpha \in [0,1]$, if, for every instance I satisfying Assumption 1,

$$\mathbb{E}[\textsc{Profit}[\Pi; I]] \geq \alpha\, \mathbb{E}[\textsc{Profit}[\textsc{Opt}; I]]\,.$$

To establish that an online policy is α-competitive, we upper-bound the offline benchmark and show that the policy attains at least a α fraction of this bound. We use this approach in Section 3 to analyze our first policy.

3. LP-Based Reservation Policy

A simple way to use the surprise-bag channel is to commit, before the selling horizon begins, to a total number of surprise-bag orders and to choose which product units to reserve for their eventual fulfillment. Neither decision is revised as regular demand unfolds. We refer to the resulting policy as the *LP-based reservation policy*. In this section, we first formulate a planning problem for the reservation, then use its solution to construct a feasible policy with the same expected profit, and finally show that it achieves the best possible competitive guarantee among all feasible online policies. To simplify notation, throughout this section we suppress the dependence on I in policy values and optimization problems.

3.1. Planning the Reservation

To evaluate a reservation decision, we use the regular-only greedy policy as the baseline. Under RO, inventory can only be sold through the regular channel or remain unsold at the end of the horizon. Reserving a unit therefore gives up the expected profit that the unit would contribute under this baseline.

For each product $i \in [N]$, index its q_i units by their sales priority under RO. Thus, unit (i,k) is sold exactly when total regular sales reach at least k, and its sale probability is $p_{i,k} \triangleq \Pr[\chi_i \geq k]$. If unit (i,k) remains in the regular channel, it earns r_i when sold and incurs the waste cost w otherwise. We call its expected profit contribution under RO the *regular value*:

$$v_{i,k} \triangleq r_i p_{i,k} - w(1 - p_{i,k}).$$

Because $p_{i,k}$ is nonincreasing in k, so is $v_{i,k}$. A unit with a larger index is less likely to earn the regular reward and more likely to incur the waste cost.

The regular values provide a unit-level decomposition of the expected profit under RO. For each product i, the tail-sum identity gives $\sum_{k\in[q_i]} p_{i,k} = \mathbb{E}[\chi_i]$. Combining this identity with (2) and the definition of $v_{i,k}$ gives

$$\mathbb{E}[\textsc{Profit}_i[\text{RO}]] = \sum\nolimits_{k\in[q_i]} v_{i,k}.$$

Let $\mathcal{U} \triangleq \{(i,k) : i \in [N],\, k \in [q_i]\}$ denote the set of all product-unit pairs. Because profit under RO is additive across products, summing the product-level decomposition over i gives

$$\mathbb{E}[\textsc{Profit}[\text{RO}]] = \sum\nolimits_{(i,k)\in\mathcal{U}} v_{i,k}.$$

We now use these unit-level values to formulate the reservation plan. Rather than choosing an integral set of units immediately, we use a fractional variable $u_{i,k} \in [0,1]$ for each $(i,k) \in \mathcal{U}$. This variable will serve as the target marginal probability that the unit is reserved for surprise-bag fulfillment.

Reserving unit (i,k) changes its contribution from $v_{i,k}$ under RO to the surprise-bag reward s, so the corresponding incremental contribution is $s - v_{i,k}$. The total reservation mass cannot exceed the number of surprise-bag orders that may be accepted over the horizon, TY. Moreover, $\sum_{k\in[q_i]} u_{i,k}$ is the target expected number of units of product i reserved for surprise-bag fulfillment, so the fractional plan must satisfy the composition requirements in (1). This gives the following static linear program (LP):

$$\begin{aligned} \mathbb{E}[\textsc{Profit}[\text{RO}]] + \max_{u\in[0,1]^{\mathcal{U}}} & \sum\nolimits_{(i,k)\in\mathcal{U}} (s - v_{i,k}) u_{i,k} \\ \text{s.t.} \quad & \sum\nolimits_{(i,k)\in\mathcal{U}} u_{i,k} \le TY, \\ & \sum\nolimits_{(i,k)\in\mathcal{U}} a_{i,\ell} u_{i,k} \le b_\ell, \qquad \ell \in [L]. \end{aligned} \tag{LP_{st}}$$

The constraints in (LP_{st}) treat all units of a given product identically, while the objective coefficient $s - v_{i,k}$ is nondecreasing in k. We can therefore choose an optimal solution with the following suffix form.

LEMMA 1 **(Suffix-form optimal solution)**. *There exists an optimal solution u^* of (LP_{st}) such that, for every product $i \in [N]$, there exist an index $k_i \in [q_i]$ and $\theta_i \in [0,1]$ satisfying*

$$u^*_{i,k} = \begin{cases} 0, & k < k_i, \\ \theta_i, & k = k_i, \\ 1, & k > k_i. \end{cases}$$

Proof. Fix an optimal solution. For any product i with indices $k < k'$ such that $u_{i,k} > 0$ and $u_{i,k'} < 1$, choose the smallest such k and the largest such k'. Move $\epsilon = \min\{u_{i,k},\, 1 - u_{i,k'}\}$ units of reservation mass from (i,k) to (i,k'). All units of product i enter the constraints with the same coefficients, so feasibility is preserved. Since $s - v_{i,k}$ is nondecreasing in k, the objective does not decrease. Each exchange either sets the chosen lower-index coordinate to zero or the chosen higher-index coordinate to one. Hence, after finitely many exchanges, no such pair remains. The positive coordinates for each product then form a suffix, with at most one fractional coordinate, which gives the stated form. □

3.2. Implementing the Reservation Plan

In the absence of the soft composition constraints, (LP_{st}) admits an integral optimal solution. With these constraints, however, an optimal solution may be fractional. We therefore seek a randomized integral reservation whose marginal reservation probabilities match the fractional solution. It preserves the soft composition constraints in expectation and respects the bag-capacity limit on every realization.

We use the dependent-rounding procedure of Chekuri et al. [2010]; Gandhi et al. [2006], given in Algorithm 1. Starting from a fractional vector, the procedure repeatedly updates two fractional coordinates while preserving their sum exactly and each coordinate in conditional expectation. If one fractional coordinate remains, it is rounded by a Bernoulli draw with success probability equal to its current value.

Algorithm 1: DEPENDENT ROUNDING (DEPROUND)

Input : A vector $x \in [0,1]^J$ over a finite index set J.

Output: An integral vector $y \in \{0,1\}^J$.

Initialize $y \leftarrow x$ and fix an ordering of J.

while *y has at least two fractional coordinates* **do**

  Let i and j be the first two fractional coordinates of y.

  Let $\delta^+ \triangleq \min\{1-y_i, y_j\}$, $\delta^- \triangleq \min\{y_i, 1-y_j\}$.

  Update $(y_i, y_j) \leftarrow \begin{cases} (y_i+\delta^+, y_j-\delta^+), & \text{w.p. } \dfrac{\delta^-}{\delta^+ + \delta^-}, \\ (y_i-\delta^-, y_j+\delta^-), & \text{w.p. } \dfrac{\delta^+}{\delta^+ + \delta^-}. \end{cases}$

if *y has exactly one fractional coordinate i* **then** set $y_i \leftarrow 1$ w.p. y_i, and $y_i \leftarrow 0$ otherwise.

The procedure has the following standard properties. It preserves each coordinate in expectation while keeping the rounded total within one unit of the fractional total. The bag-capacity bound TY is an integer, so the rounded total satisfies the bound whenever the fractional total does.

LEMMA 2. *For any input $x \in [0,1]^J$,* DEPROUND *returns a random vector $y \in \{0,1\}^J$ satisfying*

$$\mathbb{E}[y_j] = x_j \quad \textit{for all } j \in J, \qquad \sum\nolimits_{j\in J} y_j \in \left\{ \left\lfloor \sum\nolimits_{j\in J} x_j \right\rfloor, \left\lceil \sum\nolimits_{j\in J} x_j \right\rceil \right\} \quad \textit{surely}.$$

The first property follows because every randomized update preserves each coordinate in conditional expectation, while the second follows because the pairwise updates preserve the total sum exactly and leave at most one fractional coordinate before the final Bernoulli rounding. A proof is given in Section EC.5.1.

Let u^* be an optimal suffix-form solution from Lemma 1, and set $u' \leftarrow \text{DEPROUND}(u^*)$. Thus, $u'_{i,k} \in \{0,1\}$ indicates whether unit (i,k) is reserved for surprise-bag fulfillment. By Lemma 2,

$$\mathbb{E}\left[u'_{i,k}\right] = u^*_{i,k}, \quad \text{for all } (i,k) \in \mathcal{U}.$$

The floor-or-ceiling guarantee in Lemma 2, together with feasibility of u^* in (LP_{st}), also ensures that

$$\sum\nolimits_{(i,k)\in\mathcal{U}} u'_{i,k} \le TY \qquad \text{surely}.$$

For each product i, let $u'_i \triangleq \sum_{k\in[q_i]} u'_{i,k}$ denote the realized number of units reserved for surprise-bag fulfillment. The LP-based reservation policy reserves these u'_i units for surprise-bag fulfillment and makes

the remaining $q_i - u'_i$ units available only in the regular channel during the selling horizon. In each period, the policy displays as much unreserved inventory as the regular display caps allow and accepts as many surprise-bag orders as the bag cap and its remaining reserved inventory allow. Since the initial reservation is at most TY, all reserved units are committed to accepted orders by the end of the horizon. The complete policy is given in Algorithm 2.

Algorithm 2: LP-based reservation policy (SR) Policy

Find an optimal suffix-form solution u^* of ($\mathrm{LP}_{\mathrm{st}}$).
Set $u' \leftarrow \textsc{DepRound}(u^*)$ and $u'_i \leftarrow \sum_{k\in[q_i]} u'_{i,k}$ for each $i \in [N]$.
Initialize $q^R_{i,0} \leftarrow q_i - u'_i$ for each $i \in [N]$, and $q^S_0 \leftarrow \sum_{i\in[N]} u'_i$.
for $t = 1, \ldots, T$ **do**
    Set $x_{i,t} \leftarrow \min\{X_i, q^R_{i,t-1}\}$ for each $i \in [N]$, and $y_t \leftarrow \min\{Y, q^S_{t-1}\}$.
    `/* Regular demand in period t is realized. */`
    Update $q^R_{i,t} \leftarrow q^R_{i,t-1} - \min\{d_{i,t}, x_{i,t}\}$ for each $i \in [N]$, and $q^S_t \leftarrow q^S_{t-1} - y_t$.
Fulfill the accepted surprise-bag orders using u'_i units of product i for each $i \in [N]$.

LEMMA 3. *The LP-based reservation policy is feasible and has expected profit equal to the optimal value of* ($\mathrm{LP}_{\mathrm{st}}$)*, i.e.,* $\mathbb{E}[\textsc{Profit}[\mathrm{SR}]] = \mathrm{LP}_{\mathrm{st}}$.

Proof. The rounding is performed before the horizon and independently of demand, and the period-t decisions in Algorithm 2 use only the current inventory balances. Each regular display is bounded by the remaining unreserved inventory and by X_i, and each bag-acceptance decision is bounded by Y. Regular sales therefore leave the reserved units available for fulfillment, and the acceptance rule commits all reserved units by the end of the horizon. Thus every accepted order is fulfilled, and, for every product $i \in [N]$,

$$z_i = u'_i = \sum\nolimits_{k\in[q_i]} u'_{i,k}.$$

By marginal preservation in Lemma 2, $\mathbb{E}\big[u'_{i,k}\big] = u^*_{i,k}$ for every $(i,k) \in \mathcal{U}$. Thus, by linearity of expectation,

$$\mathbb{E}[z_i] = \sum\nolimits_{k\in[q_i]} \mathbb{E}\big[u'_{i,k}\big] = \sum\nolimits_{k\in[q_i]} u^*_{i,k}.$$

Feasibility of u^* in ($\mathrm{LP}_{\mathrm{st}}$) then implies that the soft composition constraints are satisfied.

We next consider profit. Since u^* is suffix-form, it has at most one fractional coordinate for each product. The rounding leaves all integral coordinates unchanged, so every realized reservation remains suffix-form. Conditional on a realized u', each reserved unit earns s. The suffix form leaves a prefix of unreserved

units for each product. These units are sold greedily in the same order as under RO, so their expected contributions are their regular values $v_{i,k}$. Hence

$$\mathbb{E}[\textsc{Profit}[\mathrm{SR}] \mid u'] = \sum\nolimits_{(i,k)\in\mathcal{U}} \left[s u'_{i,k} + v_{i,k}(1-u'_{i,k})\right].$$

Taking expectation over the rounding and applying $\mathbb{E}\left[u'_{i,k}\right] = u^*_{i,k}$ again yields $\mathrm{LP}_{\mathrm{st}}$. □

The timing of surprise-bag acceptance does not affect profit once the reservation is fixed, since any acceptance schedule that exhausts the reserved capacity and respects the per-period bag-offer cap yields the same profit.

3.3. Competitive Performance

We now evaluate the LP-based reservation policy against the offline optimal benchmark and show that its competitive guarantee is best possible among all online policies.

THEOREM 1 **(Optimal competitive ratio)**. *Let $\kappa \triangleq w/s \in [0,1]$. Under productwise regular-channel profitability (Assumption 1), the LP-based reservation policy achieves competitive ratio $\frac{1}{2+\kappa}$. In particular, it is $1/3$-competitive over all instances satisfying Assumption 1.*

Moreover, this guarantee is optimal for every $\kappa \in [0,1]$: for every $\varepsilon > 0$, there exists an instance satisfying Assumption 1 and $w/s = \kappa$ such that every feasible online policy earns at most $\frac{1}{2+\kappa} + \varepsilon$ times the offline benchmark.

Proof sketch. For the positive result, we first derive an LP upper bound on the offline optimal benchmark. On every demand realization, the regular-only greedy policy makes at least as many regular sales of each product as the offline optimal policy. Relative to this baseline, assigning a unit to a surprise bag yields at most $s+w$: this gain occurs when the unit would otherwise remain unsold, whereas assigning a unit that would sell regularly changes its contribution by $s - r_i < 0$. This yields a uniform-gain LP, denoted $\mathrm{LP}_{\mathrm{ug}}$, with the same feasible region as $\mathrm{LP}_{\mathrm{st}}$ but with every unit-level reservation gain $s - v_{i,k}$ replaced by the common upper bound $s+w$. Hence,

$$\mathbb{E}[\textsc{Profit}[\textsc{Opt}]] \le \mathrm{LP}_{\mathrm{ug}}.$$

We compare $\mathrm{LP}_{\mathrm{ug}}$ with $\mathrm{LP}_{\mathrm{st}}$ by dual fitting. Scaling an optimal solution to the dual of $\mathrm{LP}_{\mathrm{st}}$ by $1+\kappa$ yields a feasible solution to the dual of $\mathrm{LP}_{\mathrm{ug}}$. A product-by-product bound on the resulting dual objective, using Assumption 1 to control the regular-only contribution, gives

$$\mathrm{LP}_{\mathrm{ug}} \le (2+\kappa)\mathrm{LP}_{\mathrm{st}}.$$

Combining these inequalities with Lemma 3, which establishes that the expected profit of the LP-based reservation policy equals $\mathrm{LP}_{\mathrm{st}}$, proves the stated competitive ratio. The full argument for this part is given in Section EC.5.2.

For the matching impossibility result, consider an instance with one product, one period, one unit, no composition constraints, and $X_1 = Y = 1$. Let

$$s = 1, \qquad w = \kappa, \qquad \delta = \frac{1+\kappa}{r+\kappa},$$

where $r > 1$, and let regular demand equal one with probability δ and zero otherwise. If a surprise-bag order is accepted, it must be fulfilled. Hence, a feasible online policy cannot accept the bag order and expose the unit to regular demand simultaneously. The only undominated online decisions are therefore to accept the surprise-bag order or leave the unit in the regular channel. Their expected profits are, respectively,

$$1 \qquad \text{and} \qquad \delta r - (1-\delta)\kappa = 1.$$

Thus, every randomized online policy earns expected profit at most 1. The offline optimal policy observes regular demand before making this decision: it chooses the regular channel when demand is one and the surprise-bag channel otherwise, so the offline benchmark is

$$\delta r + (1-\delta) \longrightarrow 2 + \kappa \qquad \text{as } r \to \infty.$$

Hence the competitive ratio approaches $1/(2+\kappa)$. The construction also satisfies the productwise regular-channel profitability assumption because the regular-only greedy policy has expected profit 1. □

REMARK 1 (m-UNIT BAGS). Some surprise-bag offerings prescribe a fixed number of physical units per bag, so that each bag must contain exactly m units. While the aggregate composition model remains unchanged, every realized reservation must then contain a multiple of m units. This divisibility requirement must hold in every realization, whereas a direct rescaling of the one-unit reservation LP controls the marginal reservation probabilities of individual units but does not enforce this requirement pathwise. The branch-cardinality formulation developed in Section EC.4 enforces this pathwise requirement while preserving the aggregate composition model, thereby extending the static policy and its competitive analysis to m-unit bags.

4. The Adaptive Framework and Policy Dominance

The LP-based reservation policy achieves the best possible worst-case guarantee while committing to a single reservation plan before the horizon begins. The optimality of this guarantee, however, does not imply that an ex-ante commitment is best for every instance: realized regular demand changes the inventory available for future decisions, so revising future reservations may improve expected profit.

This section develops the adaptive reservation policy, which re-solves the static planning problem at the beginning of each period for the remaining horizon. It implements the resulting current-period decisions while retaining a feasible plan for the remaining periods. We show that this adaptive policy weakly dominates its static counterpart in expected profit on every instance.

4.1. Remaining-Horizon Planning

The adaptive policy re-solves the static planning problem at the beginning of each period over the remaining horizon. The corresponding residual state consists of the inventory still available for future decisions, the conditional soft budgets on the realized branch, and the conditional distribution of remaining regular demand given the observed demand history. The inventory component excludes units already sold through the regular channel or committed to accepted surprise-bag orders.

For each period t, let $\boldsymbol{d}_{1:t-1}$ denote the regular demand vectors observed in periods $1,\dots,t-1$, with $\boldsymbol{d}_{1:0}$ denoting the empty history. Let $\mathcal{D}^{(t)}(\cdot \mid \boldsymbol{d}_{1:t-1})$ denote the conditional distribution of $\{d_{i,\tau}\}_{i\in[N],\,\tau=t,\dots,T}$ given the observed history $\boldsymbol{d}_{1:t-1}$.

DEFINITION 2 (RESIDUAL STATE). At the beginning of period t, the residual state is

$$S^{(t-1)} \triangleq \left(\{\tilde{q}_i^{(t-1)}\}_{i\in[N]}, \{\tilde{b}_\ell^{(t-1)}\}_{\ell\in[L]},\ \mathcal{D}^{(t)}(\cdot \mid \boldsymbol{d}_{1:t-1})\right),$$

where $\tilde{q}_i^{(t-1)}$ is the uncommitted residual inventory of product i available for re-optimization, and $\tilde{b}_\ell^{(t-1)}$ is the row-ℓ residual soft budget on the current realized branch. Initially, $\tilde{q}_i^{(0)} \triangleq q_i$, $\tilde{b}_\ell^{(0)} \triangleq b_\ell$, and $\mathcal{D}^{(1)}(\cdot \mid \boldsymbol{d}_{1:0}) = \mathcal{D}$.

The quantities $\tilde{b}_\ell^{(t-1)}$ are not pathwise remaining budgets. They specify the future row-ℓ usage allowed in conditional expectation on the current branch. In particular, a residual soft budget may be negative even though the original right-hand side b_ℓ is nonnegative. Later we show that every state reached by the policy nevertheless carries an explicit feasible continuation.

Given the residual state $S^{(t-1)}$, re-index the units in the residual inventory by their remaining regular-sales priority and define

$$\mathcal{U}^{(t)} \triangleq \{(i,k) : i\in[N],\ k\in[\tilde{q}_i^{(t-1)}]\}.$$

As in Section 3, lower-indexed residual units of each product are those that the regular-only greedy policy would sell first. For each $(i,k)\in\mathcal{U}^{(t)}$, let $p_{i,k}^{(t)}$ denote the probability, under conditional distribution $\mathcal{D}^{(t)}(\cdot \mid \boldsymbol{d}_{1:t-1})$, that the regular-only greedy policy sells at least k units of product i over periods $t,\dots,T$ starting from the residual inventory. The residual regular value is

$$v_{i,k}^{(t)} \triangleq r_i p_{i,k}^{(t)} - w\big(1-p_{i,k}^{(t)}\big).$$

These residual unit values yield the remaining-horizon planning problem. At a residual state $S^{(t-1)}$, the residual LP is the static LP applied to the residual inventory, the remaining bag-offer capacity, and the residual soft budgets:

$$\begin{aligned}
&\sum\nolimits_{(i,k)\in\mathcal{U}^{(t)}} v_{i,k}^{(t)} + \max_{u\in[0,1]^{\mathcal{U}^{(t)}}} \sum\nolimits_{(i,k)\in\mathcal{U}^{(t)}} (s - v_{i,k}^{(t)})u_{i,k} \\
&\qquad\text{s.t.}\quad \sum\nolimits_{(i,k)\in\mathcal{U}^{(t)}} u_{i,k} \le (T-t+1)Y, \qquad\qquad (\mathrm{LP}_t(S^{(t-1)})) \\
&\qquad\qquad \sum\nolimits_{(i,k)\in\mathcal{U}^{(t)}} a_{i,\ell}u_{i,k} \le \tilde{b}_\ell^{(t-1)}, \qquad \ell\in[L].
\end{aligned}$$

As in the static LP, $u_{i,k}$ represents the mass of residual unit (i,k) reserved for surprise-bag fulfillment over the remaining horizon. When $t=1$, the residual state is the original instance, so $\mathrm{LP}_1(S^{(0)}) = \mathrm{LP}_{\mathrm{st}}$.

Since the residual regular values $v_{i,k}^{(t)}$ remain nonincreasing in k, Lemma 1 applies to ($\mathrm{LP}_t(S^{(t-1)})$). Hence, the residual LP admits a suffix-form optimal solution, which we denote by $u^{*(t)} = \{u_{i,k}^{*(t)}\}_{(i,k)\in\mathcal{U}^{(t)}}$.

4.2. Rounding to a Remaining-Horizon Schedule

This subsection formalizes the rounded schedules used by the adaptive reservation policy. We first specify the properties required of such schedules, and then give a balanced schedule construction that satisfies them.

DEFINITION 3 (ADMISSIBLE ROUNDED SCHEDULE). Fix a residual state $S^{(t-1)}$ and a feasible suffix-form solution $u^{(t)} \in [0,1]^{\mathcal{U}^{(t)}}$ of $\mathrm{LP}_t(S^{(t-1)})$. A rounded remaining-horizon schedule for $u^{(t)}$ is admissible if it consists of integer product-period counts $z_{i,\tau}^{(t)} \in \mathbb{N}, \qquad i\in[N], \quad \tau\in\{t,\dots,T\}$, with product totals $z_{i,\cdot}^{(t)} \triangleq \sum_{\tau=t}^{T} z_{i,\tau}^{(t)}$. The schedule is interpreted by the canonical suffix rule: for each product i, the schedule selects the last $z_{i,\cdot}^{(t)}$ residual units of product i, and earlier periods receive lower-indexed units within this selected suffix. Let $I_{i,k}^{(t)}$ be the induced indicator that residual unit (i,k) is selected by this rule. The counts are admissible if they satisfy the following properties:

(a) Each remaining period capacity is respected: $\sum_{i\in[N]} z_{i,\tau}^{(t)} \le Y, \qquad \tau = t,\dots,T.$

(b) Product totals fit the residual inventory: $z_{i,\cdot}^{(t)} \le \tilde{q}_i^{(t-1)}, \qquad i\in[N].$

(c) Unit-level marginals are preserved: $\mathbb{E}\left[I_{i,k}^{(t)} \,\middle|\, S^{(t-1)}\right] = u_{i,k}^{(t)}, \qquad (i,k)\in\mathcal{U}^{(t)}.$

(d) Every unexecuted tail is suffix-compatible: for every cutoff period $s\in\{t,\dots,T\}$ and product i, the units assigned to periods $s,\dots,T$ are the last $\sum_{\tau=s}^{T} z_{i,\tau}^{(t)}$ residual units of that product and hence form a suffix.

The four requirements separate the roles needed later. Requirement (a) keeps both current and future bag offers within the per-period bag-offer cap. Requirement (b) ensures that the selected suffixes are physically available. Requirement (c) preserves the residual-LP reservation probabilities in conditional expectation, which is used for both soft-budget bookkeeping and the one-step residual-LP value comparison. Requirement (d) allows both the current schedule and every unexecuted tail to be evaluated using the same unit-value accounting as in the static policy. We now give one such schedule rounding.

Balanced schedule rounding. Given a feasible suffix-form solution $u^{(t)}$, remaining periods $t,\dots,T$, and per-period bag-offer cap Y, let $u_{i,\cdot}^{(t)} = \sum_{k\in[\tilde{q}_i^{(t-1)}]} u_{i,k}^{(t)}$ and write $u_{i,\cdot}^{(t)} = \lfloor u_{i,\cdot}^{(t)} \rfloor + f_i^{(t)}$ with $f_i^{(t)} \in [0,1)$. Apply DEPROUND (Algorithm 1) to $\{f_i^{(t)}\}_{i\in[N]}$, let $\hat{f}_i^{(t)} \in \{0,1\}$ be the rounded output, and set $z_{i,\cdot}^{(t)} = \lfloor u_{i,\cdot}^{(t)} \rfloor + \hat{f}_i^{(t)}$. Let $z_{\cdot,\cdot}^{(t)} = \sum_{i\in[N]} z_{i,\cdot}^{(t)}$. Dependent rounding preserves the product-level marginals and makes this total either the floor or the ceiling of the fractional reservation mass. Because the residual capacity bound $(T-t+1)Y$ is integral, the rounded total fits within the remaining horizon.

Next construct balanced period loads. Initialize $m_t^{(t)} = z_{\cdot,\cdot}^{(t)}$. For $\tau = t,\dots,T$, set

$$z_{\cdot,\tau}^{(t)} \triangleq \left\lceil \frac{m_\tau^{(t)}}{T-\tau+1} \right\rceil,$$

and update $m_{\tau+1}^{(t)} = m_\tau^{(t)} - z_{\cdot,\tau}^{(t)}$. Feasibility of the rounded total implies $0 \le m_\tau^{(t)} \le (T-\tau+1)Y$ throughout the construction, so every period load respects the bag-offer cap and the remaining units fit within the remaining periods. Finally, fill the $z_{\cdot,\tau}^{(t)}$ slots in each period τ with product labels so that product i appears exactly $z_{i,\cdot}^{(t)}$ times across the remaining horizon. Fix any deterministic interleaving rule before the horizon; the results below hold for every such rule. Set $z_{i,\tau}^{(t)}$ equal to the number of product i labels placed in period τ, and interpret the resulting counts by the canonical suffix rule.

LEMMA 4 **(Existence of an admissible rounded schedule).** *Fix a residual state $S^{(t-1)}$ and a feasible suffix-form solution $u^{(t)} \in [0,1]^{\mathcal{U}^{(t)}}$ of $LP_t(S^{(t-1)})$. The construction above produces an admissible rounded remaining-horizon schedule for $u^{(t)}$.*

Proof sketch. Dependent rounding gives the required product totals and preserves the unit-level marginals under the canonical suffix rule. Its floor–ceiling total remains no larger than the remaining bag capacity, so the balanced load construction packs the rounded mass into periods $t,\dots,T$ while respecting (a). The product-label interleaving then realizes the prescribed product totals, which gives (b); the dependent-rounding marginals give (c); and the canonical within-product ordering gives (d). The full verification is deferred to Section EC.6.1. □

4.3. The Adaptive Reservation Policy

We now define the adaptive reservation policy. In each period it takes three steps: solve the residual LP, round its solution into a feasible remaining-horizon schedule, and execute only the schedule's current component. To keep the schedule executable, the policy withholds the selected suffix from regular sales in the current period. Units in its current component are protected for that period's bag offers, while the unexecuted tail remains provisional. After regular demand is realized, the tail is pooled with the surviving uncommitted inventory and may be reallocated when the policy re-solves; accepted bags are fulfilled at the horizon using the units protected by the realized current-period components.

For a rounded schedule $Z^{(t)}$, the canonical suffix rule induces unit-level reservation indicators: $I_{i,k}^{(t)} \triangleq \mathbb{1}\left\{k > \tilde{q}_i^{(t-1)} - z_{i,\cdot}^{(t)}\right\}$ for every $(i,k) \in \mathcal{U}^{(t)}$. Thus $I_{i,k}^{(t)} = 1$ means that unit (i,k) is temporarily withheld from the regular channel in period t. Aggregating over units of the same product gives the number of product-i units withheld from the regular channel: $z_{i,\cdot}^{(t)} \triangleq \sum_{\tau=t}^{T} z_{i,\tau}^{(t)}$. The remaining residual units are available to the regular channel. Hence the period-t regular display decision is $x_{i,t} \triangleq \min\left\{X_i, \tilde{q}_i^{(t-1)} - z_{i,\cdot}^{(t)}\right\}$ for every product $i \in [N]$. Only the schedule's current component determines the period-t bag offers. Thus, the period-t bag offer is $y_t \triangleq \sum_{i\in[N]} z_{i,t}^{(t)}$.

After regular demand is realized, we update residual inventory by removing the units protected for current bag offers and the units sold through the regular channel, $\tilde{q}_i^{(t)} \triangleq \tilde{q}_i^{(t-1)} - z_{i,t}^{(t)} - \min\{d_{i,t}, x_{i,t}\}$ for every product $i \in [N]$. The difficulty arises from rounding. The residual LP controls the fractional schedule's row usage, and marginal-preserving rounding preserves this control only in conditional expectation. A realized

rounded schedule can therefore have total current-and-tail row usage exceeding the current residual budget. After its current component is executed, subtracting only the realized current-period usage from the current residual budget need not leave enough budget for the realized tail. We instead define the next budget as the tail's realized usage plus the conditional slack left by the residual LP. We instead define the next budget as the tail's realized usage plus the conditional slack left by the residual LP. For row ℓ, define the current-period soft usage by $\sigma^{(t)}_{\ell,t}(Z^{(t)}) \triangleq \sum_{i\in[N]} a_{i,\ell} z^{(t)}_{i,t}$, and the scheduled-tail soft usage by $\sigma^{(t)}_{\ell,>t}(Z^{(t)}) \triangleq \sum_{\tau=t+1}^{T} \sum_{i\in[N]} a_{i,\ell} z^{(t)}_{i,\tau}$. The conditional slack is $\mathsf{Slack}^{(t)}_{\ell} \triangleq \tilde{b}^{(t-1)}_{\ell} - \mathbb{E}\left[\sigma^{(t)}_{\ell,t}(Z^{(t)}) + \sigma^{(t)}_{\ell,>t}(Z^{(t)}) \,\middle|\, S^{(t-1)}\right]$. The next residual soft budget adds this slack to the tail's realized usage: $\tilde{b}^{(t)}_{\ell} \triangleq \sigma^{(t)}_{\ell,>t}(Z^{(t)}) + \mathsf{Slack}^{(t)}_{\ell}$ for every $\ell \in [L]$. The expectation in the slack is directly computable from the residual-LP solution: marginal preservation gives

$$\mathbb{E}\left[\sigma^{(t)}_{\ell,t}(Z^{(t)}) + \sigma^{(t)}_{\ell,>t}(Z^{(t)}) \,\middle|\, S^{(t-1)}\right] = \sum_{(i,k)\in\mathcal{U}^{(t)}} a_{i,\ell} u^{*(t)}_{i,k}.$$

Thus no enumeration of rounded schedules is required. Whenever the current residual LP is feasible, this identity implies $\mathsf{Slack}^{(t)}_{\ell} \geq 0$ and hence

$$\sigma^{(t)}_{\ell,>t}(Z^{(t)}) \leq \tilde{b}^{(t)}_{\ell}, \tag{3}$$

$$\mathbb{E}\left[\sigma^{(t)}_{\ell,t}(Z^{(t)}) + \tilde{b}^{(t)}_{\ell} \,\middle|\, S^{(t-1)}\right] = \tilde{b}^{(t-1)}_{\ell}. \tag{4}$$

The first relation makes the realized tail feasible for the next soft row; the second ensures that expected current usage plus the expected next residual budget equals the current residual budget. Because the coefficients $a_{i,\ell}$ may be signed, the resulting residual budget need not be nonnegative. The tail witness, rather than the zero plan, certifies feasibility of the next residual LP. The formal definition of the adaptive reservation policy is given in Algorithm 3.

Feasibility analysis. We now verify that the adaptive reservation policy is feasible for the original problem.

LEMMA 5 **(Tail feasibility witness).** *Fix a period $t < T$ and a residual state $S^{(t-1)}$ at which the residual LP is feasible. Let $Z^{(t)}$ be the admissible rounded remaining-horizon schedule drawn from $u^{*(t)}$. After the period-t bag commitments and regular sales are realized, the unexecuted tail of $Z^{(t)}$ induces a feasible integral solution to the residual LP at the next residual state $S^{(t)}$.*

Proof sketch. The canonical suffix rule makes the unexecuted tail physically disjoint from the period-t bag commitments and protected from period-t regular sales. Properties (a) and (b) give the remaining hard-capacity and inventory feasibility. Marginal preservation and residual-LP feasibility make the common conditional slack nonnegative, so (3) gives soft-row feasibility even when the next residual budget is negative. See Section EC.6.2 for the full proof. □

PROPOSITION 1 **(Feasibility of adaptive reservation policy).** *The adaptive reservation policy is feasible for every sample path and satisfies the bag composition constraints (1) in expectation.*

Algorithm 3: Adaptive Reservation (AR) Policy

Initialize $S^{(0)} = \big(\{q_i\}_{i\in[N]}, \{b_\ell\}_{\ell\in[L]}, \mathcal{D}\big)$.

for $t = 1, \dots, T$ **do**

  Solve the residual LP in $\mathrm{LP}_t(S^{(t-1)})$, and choose a suffix-form optimal solution $u^{*(t)} = \{u^{*(t)}_{i,k}\}_{(i,k)\in\mathcal{U}^{(t)}}$.

  Draw an admissible rounded remaining-horizon schedule $Z^{(t)}$ for $u^{*(t)}$ using Lemma 4.

  Display $x_{i,t}$ units for each product $i \in [N]$, and offer y_t bags, as determined by the rounded schedule $Z^{(t)}$.

  `/* Regular demand in period t is realized. */`

  Update the residual inventory $\{\tilde{q}_i^{(t)}\}_{i\in[N]}$ and the residual soft budget $\{\tilde{b}_\ell^{(t)}\}_{\ell\in[L]}$, and form the next residual state $S^{(t)} = \Big(\{\tilde{q}_i^{(t)}\}_{i\in[N]}, \{\tilde{b}_\ell^{(t)}\}_{\ell\in[L]}, \mathcal{D}^{(t+1)}(\cdot \mid \boldsymbol{d}_{1:t})\Big)$.

Fulfill all bag offers over the horizon using $z_i \triangleq \sum_{t\in[T]} z_{i,t}^{(t)}$ units for each product $i \in [N]$.

Proof sketch. The zero plan makes the initial residual LP feasible, and Lemma 5 inductively keeps every later residual LP feasible. The hard constraints follow directly from the policy definitions and admissibility: product availability bounds regular displays, period capacity bounds current bag offers, and the units protected for those offers are removed from residual inventory and remain available for final fulfillment. For the soft constraints, (4) telescopes in expectation; the terminal residual soft budget is nonnegative because the last tail is empty. The full argument is in Section EC.6.3. □

4.4. Value of Re-solving

At the next period, the policy could simply continue with the unexecuted tail of its previous schedule. Because that tail remains feasible, re-solving the residual LP can only improve upon this fallback in conditional expected value. This observation yields a one-step comparison that telescopes over the horizon. We formalize it using the optimal residual-LP value $\mathrm{LP}_t(S^{(t-1)})$ and the terminal convention $\mathrm{LP}_{T+1}(S^{(T)}) \triangleq 0$. Let P_t be the policy's realized period-t contribution, with terminal waste charged in period T, so that $\sum_{t=1}^{T} P_t = \textsc{Profit}[\mathrm{AR}; I]$ pathwise.

We use the same unit-value accounting as for the static policy: conditional on a residual state, a feasible suffix reservation earns the residual regular-only value plus $s - v_{i,k}^{(t)}$ for each reserved unit. This identity remains valid with negative residual budgets because it depends only on suffix structure and the conditional distribution of remaining demand carried in the residual state.

Lemma 6 **(One-step residual-LP value comparison)**. *For every period t and every residual state $S^{(t-1)}$ reached by the adaptive reservation policy,*

$$\mathbb{E}\big[P_t + LP_{t+1}(S^{(t)}) \,\big|\, S^{(t-1)}\big] \geq LP_t(S^{(t-1)}).$$

Proof sketch. Fix the residual state and compare the algorithm to the fixed continuation that executes the current component of $Z^{(t)}$ and then follows its unexecuted tail. The canonical suffix rule and (a)–(b) give a feasible suffix reservation in the residual instance. The unit-value accounting above and expectation over the rounding, together with (c), identify this fixed-continuation value with the current residual-LP value. The terminal case follows directly from the empty tail and the convention $\mathrm{LP}_{T+1}(S^{(T)}) = 0$. For $t < T$, Lemma 5 makes the unexecuted tail feasible for the next residual LP, so the next residual-LP value is at least the fixed tail-continuation value. Combining the two comparisons proves the lemma; see Section EC.6.4. □

THEOREM 2 **(Adaptive weak dominance)**. *For every instance I,*

$$\mathbb{E}[\textsc{Profit}[\mathrm{AR}; I]] \geq \mathbb{E}[\textsc{Profit}[\mathrm{SR}; I]]\,.$$

Thus the adaptive reservation policy weakly dominates the LP-based reservation policy.

Proof. Taking expectations in Lemma 6 gives

$$\mathbb{E}[P_t] \geq \mathbb{E}\left[\mathrm{LP}_t(S^{(t-1)})\right] - \mathbb{E}\left[\mathrm{LP}_{t+1}(S^{(t)})\right].$$

Summing over $t = 1, \ldots, T$ telescopes the residual-LP values. Using the terminal convention, the pathwise profit decomposition, and $\mathrm{LP}_1(S^{(0)}) = \mathrm{LP}_{\mathrm{st}}$, we obtain

$$\mathbb{E}[\textsc{Profit}[\mathrm{AR}; I]] \geq \mathrm{LP}_{\mathrm{st}} = \mathbb{E}[\textsc{Profit}[\mathrm{SR}; I]]\,,$$

where the equality is Lemma 3. □

COROLLARY 1 **(Competitive ratio of adaptive reservation policy)**. *The adaptive reservation policy achieves competitive ratio $\frac{1}{2+\kappa}$ under productwise regular-channel profitability (Assumption 1).*

Proof. The result follows by combining Theorems 1 and 2. □

5. Value of Deferred Assignment

Our baseline model permits *deferred assignment* (DA): accepting a surprise-bag order creates a fulfillment obligation without immediately designating the product that will fulfill it. To isolate the value of this flexibility, we compare DA with the *immediate-assignment* (IA) counterfactual, which irrevocably assigns one unit of a specified product to each accepted bag before current-period regular demand is observed. DA can replicate any IA policy by recording each of its product commitments and following its acceptance and display decisions. Thus, DA weakly dominates IA on every instance. Writing $V^{\mathrm{DA}}(I)$ and $V^{\mathrm{IA}}(I)$ for the respective optimal expected profits,

$$V^{\mathrm{DA}}(I) \geq V^{\mathrm{IA}}(I).$$

The substantive question is when the inequality is strict. This flexibility matters when accepting an early bag is attractive but it is uncertain which product is more valuable to retain for regular sales. DA can use

later demand information to fulfill the bag with the product that proves less valuable to retain for regular sales, potentially yielding higher profit than IA.

Weak dominance need not yield a strict profit improvement. The following two examples illustrate distinct reasons why DA and IA can have the same optimal expected profit: waiting can make DA unnecessary, or reassignment can make no difference. They differ only in their regular-demand processes. Each instance has $N = 2$ products and $T = 2$ periods, with $q_i = 1$ unit of inventory and $X_i = 1$ regular-display capacity for each product i. The per-period surprise-bag cap is $Y = 1$. The regular reward is $r_i = 2$ for each product i, the surprise-bag reward is $s = 1$, and the waste cost is $w = 0$.

EXAMPLE 1 (WAITING UNTIL THE FINAL PERIOD). Regular demand satisfies

$$(d_{1,1}, d_{2,1}) = \begin{cases} (1,0), & \text{with probability } 1/2, \\ (0,1), & \text{with probability } 1/2. \end{cases} \qquad (d_{1,2}, d_{2,2}) = (0,0).$$

At first sight, DA appears more valuable. After accepting a period-1 surprise-bag order, a DA policy can display both products in the regular channel, sell one unit there, and use the remaining unit to fulfill the order, earning $2+1=3$. Under IA, accepting the same order requires assigning it to a product before regular demand is observed, leaving only the other product available for regular sale. The policy earns 3 if demand targets the product left available and 2 otherwise, when the other product is used for the period-1 bag and the unsold product can be used for a period-2 bag. Its expected profit is therefore $\frac{1}{2}(2+1)+\frac{1}{2}(2)=\frac{5}{2}<3$.

An IA policy can instead earn 3 by declining the period-1 surprise-bag order. It displays both products in the regular channel in period 1, so exactly one product sells, and then assigns the remaining unit to a period-2 surprise-bag order. No policy under either regime can earn more: at most one of the two units can earn the regular reward of 2, and every other unit earns at most the surprise-bag reward of 1. The optimal expected profits under DA and IA are therefore equal: waiting until the final period makes DA unnecessary. □

EXAMPLE 2 (WEAK REGULAR DEMAND). Regular demand satisfies

$$d_{i,1} \sim \text{Bernoulli}\left(\frac{1}{4}\right) \quad \text{independently for } i = 1,2, \qquad d_{i,2} = 0 \quad \text{for } i = 1,2.$$

Before period-1 demand is observed, the retailer decides under either regime whether to accept the period-1 surprise-bag order. We consider acceptance and rejection in turn.

If the order is accepted, the source of equality differs from Example 1: even under DA, the retailer must withhold one unit from regular display to ensure fulfillment if both products have demand. Displaying both would leave no unit for the bag with probability $1/16$. Thus, DA and IA must each withhold one unit from period-1 regular demand. DA can still choose which product remaining after regular sales fulfills the order, but this flexibility does not improve expected profit.

If the order is declined, both regimes optimally display both products in period-1 and offer a period-2 bag whenever inventory remains. Because period-2 regular demand is zero, assignment timing cannot affect this policy.

Under either regime, direct calculation gives a conditional optimal expected profit of $\frac{9}{4}$ upon accepting the period-1 surprise-bag order and $\frac{31}{16}$ upon rejecting it. Since $\frac{9}{4} > \frac{31}{16}$, acceptance is optimal under both regimes. Thus, unlike Example 1, IA does not recover by waiting; reassignment instead makes no difference because regular demand is too weak. The optimal expected profits are therefore both $\frac{9}{4}$. □

5.1. When Deferred Assignment Creates Strict Value

The two examples show that DA alone need not improve profit. They also suggest that a strict DA–IA gap can depend on both a limited per-period bag-offer cap and regular-demand uncertainty. We therefore identify sufficient conditions under which DA has strictly higher expected profit. We focus on symmetric instances that remove ex ante product heterogeneity and isolate the value of DA. There are $N \geq 2$ products and $T \geq 2$ periods. Each product $i \in [N]$ has display cap $X_i \triangleq X \geq 1$, initial inventory $q_i \triangleq q \triangleq TX$, and regular reward $r_i \triangleq r$. The surprise-bag reward and waste cost are s and w, respectively. In this symmetric setting, $r > s > 0$ and $s \geq w \geq 0$. In every period, at most $Y \geq 1$ surprise-bag orders can be accepted, and there are no surprise-bag composition constraints.

Regular demands are i.i.d. across products and periods, with $d_{i,t} \sim \mathrm{Uniform}\{0, 1, \ldots, D\}$ for some $D \in \mathbb{N}$:

$$\Pr[d_{i,t} = d] = \frac{1}{D+1}, \qquad d \in \{0, 1, \ldots, D\}.$$

Let $J \triangleq \big(N,\ T,\ X,\ Y,\ r,\ s,\ w\big)$ collect the remaining primitives, and write $I \triangleq \big(J, D\big)$. We call I a *symmetric uniform-demand instance*. A policy is *IA-feasible* (*DA-feasible*) if it is feasible under IA (DA), and *IA-optimal* (*DA-optimal*) if it maximizes expected profit among IA-feasible (DA-feasible) policies.

Under a limited per-period bag-offer cap, the theorem below gives ordered demand thresholds such that every integer demand bound strictly between them yields a strict DA advantage. Its upper threshold uses a surplus polynomial that aggregates favorable inventory-surplus histories in which an additional early bag can be fulfilled without changing subsequent decisions inherited from the corresponding IA policy. For $\eta > 0$, define

$$S(\eta) := \sum_{m=1}^{T-1} (1+\eta)^{N(m-1)} \sum_{j=1}^{N} \binom{N}{j} \sum_{\substack{n_1,\ldots,n_X \geq 0 \\ n_1+\cdots+n_X = j \\ \sum_{r=1}^{X} r n_r > mY}} \binom{j}{n_1, \ldots, n_X} \left(\frac{\eta}{X}\right)^j. \tag{5}$$

Let $\phi(J)$ be the unique positive solution of

$$\phi(J) + \big(1+\phi(J)\big) S(\phi(J)) = \frac{r-s}{s+w}. \tag{6}$$

Define the two demand thresholds

$$D_{\mathrm{L}} := \begin{cases} X - 1, & 2Y \leq (N-1)X, \\ \dfrac{(r+w)X}{r-s} - 1, & 2Y > (N-1)X, \end{cases} \tag{7}$$

$$D_{\mathrm{U}} := \frac{X}{\phi(J)} + X - 1. \tag{8}$$

These are sufficient thresholds, not intrinsic boundaries of the region in which DA has strict value.

THEOREM 3. *Let $I = (J, D)$ be a symmetric uniform-demand instance, and let DA^* and IA^* denote arbitrary DA-optimal and IA-optimal policies, respectively. Suppose*

(i) ***Limited surprise-bag offer capacity:*** $Y \le (N-1)X$.

Then $D_{\mathrm{L}} < D_{\mathrm{U}}$. Moreover, if the demand bound D satisfies

(ii) ***Demand above lower threshold:*** $D > D_{\mathrm{L}}$*; and*

(iii) ***Demand below upper threshold:*** $D < D_{\mathrm{U}}$.

then $\mathbb{E}[\textsc{Profit}[\mathrm{DA}^*; I]] > \mathbb{E}[\textsc{Profit}[\mathrm{IA}^*; I]]$.

The theorem identifies an intermediate-demand regime with a limited per-period bag-offer cap, in which early bag acceptance is valuable but uncertainty in regular demand makes its product assignment consequential. The following limiting cases explain the role of its three conditions.

Condition (i) excludes effectively unlimited bag capacity. If bag capacity were unlimited, neither regime would need to accept bags before period T: both could use only the regular channel through period $T-1$ and wait until period T to accept bags. Under independent product demands, every unit displayed in the regular channel in period T sells with positive probability. DA must therefore withhold one unit from the regular channel for each accepted bag. IA can assign those units immediately, so the two regimes have the same optimal expected profit.

Condition (ii) excludes sufficiently weak regular demand, under which DA has little regular-sale profit to protect. At the extreme, when $w = 0$ and $D = 0$, each product's regular-only profit is zero. Thus, Assumption 1 holds with equality even though no inventory can be sold through the regular channel. DA and IA therefore sell as many units through surprise bags as inventory and the per-period bag-offer cap permit and earn the same profit.

Condition (iii) excludes arbitrarily strong regular demand. As $D \to \infty$, a regular-only policy sells all inventory with probability approaching one. Because regular sales earn more than surprise bags, the surprise-bag channel becomes redundant in the limit, and the optimal-profit advantage of DA over IA vanishes.

Together, these limiting cases explain why restrictions of these qualitative types are needed; they do not establish the theorem's exact sufficient cutoffs.

Proof sketch of Theorem 3. By Lemma EC.3, select an IA-optimal policy that displays every unit not assigned to a bag, up to the regular-display cap. Construct its DA shadow by keeping the same bag acceptances and display decisions while postponing product assignments until the end of the horizon. By Definition EC.1 and Lemma EC.4, the shadow is DA-feasible and has the same realized profit as the IA policy. It therefore suffices to construct a DA-feasible policy with strictly larger expected profit than the shadow.

Assume the limited bag-capacity condition in Condition (i). Consider the full-sales reference history, on which every regular display before the last period sells out. Along this history, the IA-optimal policy either accepts at least one bag before the last period or accepts none. We analyze these two cases separately.

If IA accepts a bag before the last period on the reference history, then the bag's product assignment is fixed before subsequent regular demand is known. From this early-acceptance state, we construct a positive-probability demand continuation on which DA can reassign the bag to another product and release the originally assigned product for an additional regular-sale opportunity. The lower demand threshold in Condition (ii) guarantees that this reassignment yields a strictly positive expected gain.

If IA accepts no bag before the last period on the reference history, then it leaves the prefinal bag-offer capacity unused along that history. We instead construct a policy that accepts one additional period-1 bag and postpones its product assignment. If IA later accepts a bag, the additional early bag can replace that later bag without reducing profit. Otherwise, on favorable demand histories, enough inventory remains to fulfill the additional bag without changing subsequent decisions of the shadow. The upper demand threshold in Condition (iii) guarantees that the expected benefit from these favorable histories outweighs the possible loss if the additional bag ultimately requires withholding one unit from regular sale in the last period.

The two cases are exhaustive. In either case, a DA-feasible policy strictly improves upon the DA shadow. Since the DA shadow has the same profit as the IA-optimal policy, it follows that $V^{\mathrm{DA}}(I) > V^{\mathrm{IA}}(I)$. □

The full policy constructions and verification are given in Section EC.7.

5.2. Exact DP Evidence on When Deferred Assignment Adds Value

Examples 1 and 2 show why DA need not create value: IA can sometimes recover by accepting a bag later, or regular demand can be too weak for the product used to fulfill a bag to matter. We use exact finite-horizon dynamic programming (DP) calculations to examine how the DA–IA gap varies with regular-demand strength and the per-period bag-offer cap, identify the operating regimes in which DA has positive value or no value, and compare these patterns with the sufficient region in Theorem 3.

Numerical setting and exact DP calculation. We consider symmetric instances with $N = 3$ products and $T = 3$ periods. Each product has display cap $X = 2$ and initial inventory $q_i = TX = 6$. A regular sale earns $r = 2$, a surprise bag earns $s = 1$, disposal has zero cost, and there are no bag-composition constraints. The choices $N = T = 3$ and $X = 2$ provide a multi-product, multi-period inventory setting while keeping the state space tractable for exact DP calculation. Setting $w = 0$ isolates assignment timing from disposal, while $r = 2$ and $s = 1$ make a regular sale worth twice a surprise bag.

For each instance, we solve exact finite-horizon DPs for IA and DA and report the relative DA gain

$$100\,\frac{V^{\mathrm{DA}} - V^{\mathrm{IA}}}{V^{\mathrm{IA}}}.$$

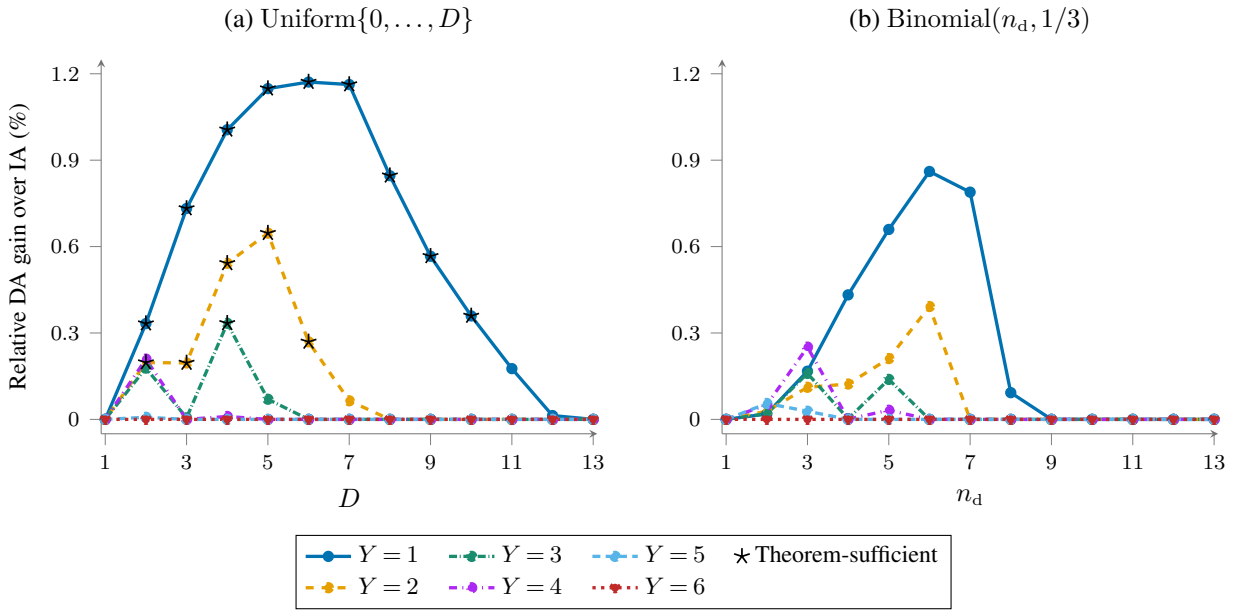


Figure 1 Exact relative DA gain over IA (%) by regular-demand strength and the per-period bag-offer cap. (a) Independent $\mathrm{Uniform}\{0,\ldots,D\}$ **product demands; stars mark the instances satisfying the sufficient conditions in Theorem 3. (b) Independent** $\mathrm{Binomial}(n_{\mathrm{d}},1/3)$ **product demands. Each curve corresponds to one** $Y\in\{1,\ldots,6\}$**. All other parameters are** $N=T=3$**,** $X=2$**,** $q_i=6$**,** $r=2$**,** $s=1$**, and** $w=0$**.**

Demand strength and the per-period bag-offer cap. Figure 1 reports the relative DA gain across per-period bag-offer caps and demand strengths under two demand families with independent product demands. Figure 1(a) uses the uniform-demand family in Theorem 3,

$$d_{i,t} \overset{\text{i.i.d.}}{\sim} \mathrm{Uniform}\{0,\ldots,D\}, \qquad D\in\{1,\ldots,13\}.$$

The starred instances satisfy the sufficient conditions in Theorem 3, independently of the DP results. Figure 1(b) changes only the marginal demand family:

$$d_{i,t} \overset{\text{i.i.d.}}{\sim} \mathrm{Binomial}(n_{\mathrm{d}},1/3), \qquad n_{\mathrm{d}}\in\{1,\ldots,13\}.$$

Thus each product-period demand is the number of purchases among n_{d} independent opportunities, each resulting in a purchase with probability $1/3$. The horizontal axes are family-specific demand-scale parameters and are not intended for pointwise cross-panel comparison.

Our numerical results in Figure 1 illustrate both the scope and the limitations of the theorem's sufficient conditions. In Figure 1(a), every theorem-certified instance has a strictly positive exact DP gain, as guaranteed by Theorem 3, but strict DA value also extends beyond the certified region. These additional positive-gain instances are consistent with the theorem imposing simple sufficient conditions rather than characterizing all instances with strict DA value.[5] Under the tightest bag-offer cap, $Y=1$, the uniform-demand gain peaks at 1.17% at an intermediate demand level.

[5] In additional computational diagnostics, we examined these off-region positive-gain instances using the two policy ideas underlying the proof: reassigning an earlier bag and accepting an additional early bag. These diagnostics find strict gains from the same

Figure 1(b) shows that strict DA value is not specific to the theorem's uniform-demand family. Under binomial demand, the gain reaches 0.86% under the tightest bag-offer cap, $Y = 1$.

Taken together, the two demand families reveal a common operating pattern. DA gains are concentrated at intermediate demand and tend to be largest under tight bag-offer caps. As later bag-offer flexibility increases, the gains generally contract. At $Y = 6$, IA and DA coincide throughout the displayed demand ranges because IA has enough later bag-offer capacity to defer acceptance and avoid costly early product commitment. At weak demand, little regular-sale profit is at stake; at strong demand, little inventory remains for product-assignment flexibility to matter. DA therefore matters most at intermediate demand when later bag-offer opportunities are limited.

6. Numerical Experiments

In this section, we complement the theoretical analysis with a case study of the dual-channel surplus-food problem. We evaluate how the proposed reservation policies perform under demand distributions calibrated from transaction data and compare them with natural heuristics and information-relaxed benchmarks.

Data and instance construction. We partnered with a Chinese catering company that operates corporate cafeterias and obtained transaction-level data from one site whose primary service is weekday lunch. The dataset records sales of ready-to-eat dishes during the lunch window 10:45–12:45 over the 27 weekdays in September 2021. Each day's window is divided into eight consecutive 15-minute periods, yielding a panel of $27 \times 8 = 216$ period-level observations. We aggregate items into 12 price categories defined by their regular price, with $r_i \in \{4, 5, \ldots, 15\}$ CNY. Mean daily sales across all categories total approximately 285 units. Because the dataset records sales rather than latent demand, it does not reveal whether sales are censored by stockouts; throughout this case study, we use observed sales as a demand proxy.

We calibrate the experimental instance from the empirical data as follows. Initial inventory $\tilde{q}_{i,0}$ for each price category is set to twice the observed average daily sales. For each product-period pair (i, t), regular demand $d_{i,t}$ is sampled from the empirical marginal distribution of observed integer sales for product i in period t.[6] Given the limited sample size, we use independent marginal resampling to construct calibrated instances. We set the bag reward to $s = 3$, use the product-independent waste cost $w = 1.5$, set the bag-offer cap to $Y = 60$, and take regular-display caps to be nonbinding. To instantiate the average-quality floor requirement in Section EC.2, we use each product's regular price as its quality score. We consider $\overline{r} \in \{0, 6, 8\}$; each positive setting requires the ratio of expected total regular price to the expected number of surprise-bag units to be at least $\overline{r}$. The setting $\overline{r} = 0$ imposes no soft constraint.

Implemented policies. We evaluate the policies and benchmarks defined in Sections 3 and 4 and Section EC.9.

ideas beyond the theorem's sufficient region, although some instances require refinements of the specific policy constructions used in the proof.

[6] If no sale is observed for product i in period t on a given date, we record the corresponding sales observation as zero.

*Regular-only greedy policy (*RO*):* It uses no demand-distribution information and never allocates inventory to the surprise-bag channel, so all sales go through the regular channel.

*Uniform-share policy (*US*):* It uses the demand distribution at time 0 and restricts LP-based reservation policy to a common bag fraction across products.

*Certainty-equivalent policy (*CE*):* It uses only mean demand information at time 0 and replaces unit-level regular values with certainty-equivalent values computed from those means.

*LP-based reservation policy (*SR*):* It uses the demand distribution at time 0 and solves the static LP; when an average-quality floor is active, that floor is included in the LP before the rounded reservation plan is executed.

*Adaptive reservation policy (*AR*):* It uses the demand distribution together with the realized residual inventory and residual soft budget, and re-solves the residual LP period by period.

*Offline optimal benchmark (*OPT*):* For the no-floor setting $\overline{r} = 0$, it uses the demand realization ex post and therefore represents the fully informed offline optimum.

*Ex-ante benchmark (*EA*):* For the active-floor settings $\overline{r} \in \{6, 8\}$, it solves the ex-ante relaxation $\mathrm{LP}_{\mathrm{ea}}$. The relaxation enforces the bag-capacity and average-quality constraints in expectation, and its optimal value upper-bounds the expected profit of every feasible online policy.

Evaluation protocol. Each policy other than the ex-ante benchmark EA is evaluated on 1,000 independently generated demand realizations from the calibrated demand model. The ex-ante benchmark is computed directly from $\mathrm{LP}_{\mathrm{ea}}$ rather than simulated. All simulated policies are evaluated on the same demand realizations to enable paired comparisons. For policies with randomized rounding, we repeat the simulation over 5 independent random seeds and report averages over both demand realizations and policy randomness.

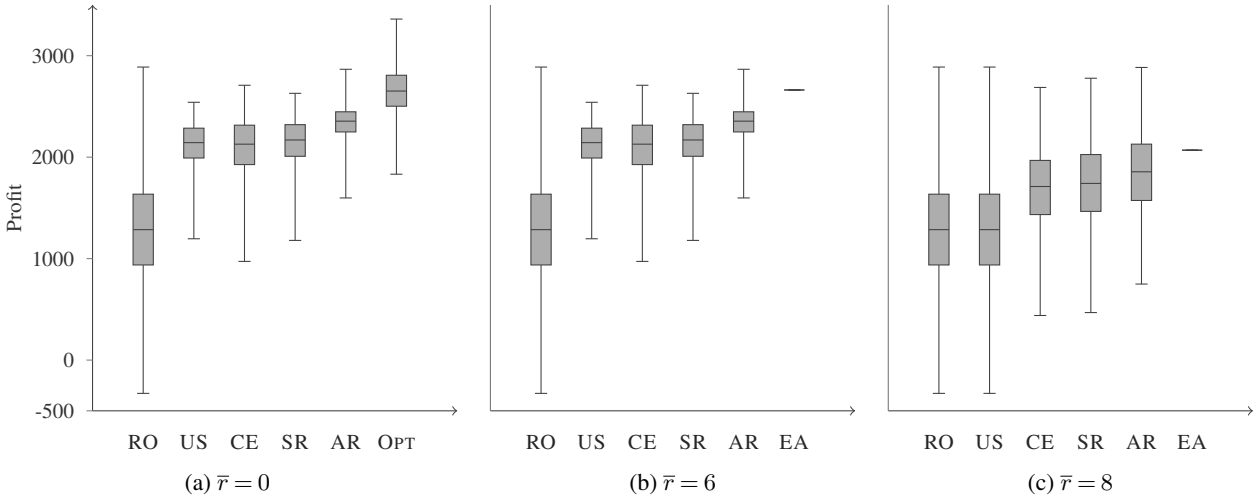
$\overline{r}$	RO	US	CE	SR	AR	OPT/EA
0	1,289.86	2,126.33	2,105.14	2,149.33	2,347.35	2,656.55
6	1,289.86	2,126.33	2,105.14	2,149.33	2,347.35	2,661.99
8	1,289.86	1,289.86	1,702.60	1,739.53	1,847.56	2,068.77

Table 1 Mean profit at $\overline{r} \in \{0, 6, 8\}$. The final column reports the relevant upper bound: OPT when $\overline{r} = 0$ and EA when $\overline{r} \in \{6, 8\}$.

Main results. Table 1 reports the mean profit at the three settings $\overline{r} \in \{0, 6, 8\}$, and Figure 2 shows the corresponding profit distributions across demand realizations. The final column reports the relevant upper bound for each setting: OPT when $\overline{r} = 0$ and EA when the average-quality floor is active.

The results show three main patterns.

First, routing inventory through the surprise-bag channel creates substantial value relative to the regular-only baseline. When the average-quality constraint is absent ($\overline{r} = 0$) or nonbinding ($\overline{r} = 6$), SR increases



(a) $\overline{r} = 0$ (b) $\overline{r} = 6$ (c) $\overline{r} = 8$

Figure 2 Profit distributions across simulated policies at $\overline{r} \in \{0, 6, 8\}$. The EA entries at $\overline{r} \in \{6, 8\}$ are deterministic upper bounds, so their boxes collapse to a line.

mean profit from 1,289.86 to 2,149.33, a 66.6% improvement over RO. AR further increases mean profit to 2,347.35, corresponding to an 82.0% improvement over RO.

Second, AR consistently improves performance. Across all three average-quality settings, AR is the best-performing policy in our implementation. Relative to SR, it improves mean profit by 9.2% when $\overline{r} \in \{0, 6\}$ and by 6.2% when $\overline{r} = 8$. This gain reflects the value of re-solving the reservation problem as regular sales update the available inventory.

Third, the binding quality floor exposes a limitation of uniform-share reservation. At $\overline{r} = 8$, the common-fraction restriction makes zero reservation optimal for US, so its mean profit equals the regular-only baseline. CE, SR, and AR can instead vary reservation quantities across products and continue to earn higher mean profits than RO.

Finally, AR remains close to the relevant benchmark upper bound. It achieves 88.4% of OPT when $\overline{r} = 0$, 88.2% of EA when $\overline{r} = 6$, and 89.3% of EA when $\overline{r} = 8$. Equivalently, its mean-profit gap relative to the benchmark is approximately 10.7%–11.8% across all three settings.

Overall, this case study shows that demand-aware reservation improves dual-channel inventory routing and that adaptive re-solving adds value across the tested average-quality settings.

7. Conclusion and Future Directions

We studied how a retailer can allocate a fixed inventory of perishable products between regular sales and surprise-bag commitments over a finite selling horizon. The model captures the tradeoff between preserving inventory for uncertain regular demand, earning guaranteed bag revenue, and incurring waste costs on unsold inventory. It also accommodates linear constraints on the expected aggregate composition of bag allocations. Unit-level regular values provide a tractable measure of the opportunity cost of reserving inventory. Using these values, an LP-based static reservation policy achieves the optimal competitive

ratio $1/(2 + w/s)$ under productwise regular-channel profitability. An adaptive policy that re-solves the remaining-horizon problem weakly dominates the static policy in expected profit while preserving feasibility. The static guarantee extends to bags containing a fixed number of units.

The analysis distinguishes the value of revising future reservations from the value of deferring product assignment for accepted bags. The proposed policies benefit from reservation decisions based on inventory values, while a separate comparison identifies conditions under which deferred assignment strictly improves optimal expected profit over immediate assignment. Simulations calibrated from cafeteria sales data further illustrate the benefit of adaptive reservation across the tested quality-floor settings. These results concern profit and fulfillment within the specified model; they do not establish the effect of surprise-bag sales on total food waste or consumer welfare.

Several extensions merit further study. First, endogenous bag composition and requirements imposed on each individual bag would connect online commitments with a packing problem at fulfillment. Unlike aggregate constraints in expectation, such requirements would require the retailer to preserve enough inventory of compatible products to assemble every accepted bag. Developing tractable policies with performance guarantees under these constraints would broaden the framework's applicability.

Second, uncertain surprise-bag demand, cancellations, and pickup failures would make both revenue and fulfillment obligations stochastic. Allowing demand to respond to prices and past bag quality would also connect intraday inventory allocation with customer behavior over repeated selling days. Finally, the current policies assume a known regular-demand distribution and observation of realized demand. Learning from limited and potentially censored sales data raises the question of how estimation error and partial demand feedback affect reservation decisions and their performance guarantees.

References

Arzum Akkaş and Vishal Gaur. OM Forum—Reducing Food Waste: An Operations Management Research Agenda. *Manufacturing & Service Operations Management (M&SOM) (INFORMS)*, 24(3):1261–1275, May 2022. ISSN 1523-4614. doi: 10.1287/msom.2021.1044. URL https://research.ebsco.com/plink/c89ea41c-43c3-3580-b563-a82ee0b5575a.

Saeed Alaei, MohammadTaghi Hajiaghayi, and Vahid Liaghat. Online prophet-inequality matching with applications to ad allocation. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, EC '12, pages 18–35, New York, NY, USA, 2012. Association for Computing Machinery. doi: 10.1145/2229012.2229018. URL https://doi.org/10.1145/2229012.2229018.

Moshe Babaioff, Jason D. Hartline, and Robert D. Kleinberg. Selling ad campaigns: online algorithms with cancellations. In *Proceedings of the 10th ACM Conference on Electronic Commerce*, EC '09, page 61–70, New York, NY, USA, 2009. Association for Computing Machinery. ISBN 9781605584584. doi: 10.1145/1566374.1566383. URL https://doi.org/10.1145/1566374.1566383.

Gérard P. Cachon and A. G. Kök. Implementation of the newsvendor model with clearance pricing: How to (and how not to) estimate a salvage value. *Manuf. Serv. Oper. Manag.*, 9:276–290, 2007. doi: 10.1287/msom.1060.0145.

Chandra Chekuri, Jan Vondrák, and Rico Zenklusen. Dependent randomized rounding via exchange properties of combinatorial structures. In *2010 IEEE 51st Annual Symposium on Foundations of Computer Science*, 2010. doi: 10.1109/FOCS.2010.60. URL http://ieeexplore.ieee.org/document/5671314/.

Alexandra Yizhuo Dong, Auyon Siddiq, and Jingwei Zhang. Empirical market design in surplus food marketplaces, 2025. URL https://ssrn.com/abstract=5648812. Available at SSRN.

Farbod Ekbatani, Yiding Feng, and Rad Niazadeh. Online resource allocation with buyback: Optimal algorithms via primal-dual. In *Proceedings of the 24th ACM Conference on Economics and Computation*, EC '23, page 583, New York, NY, USA, 2023. Association for Computing Machinery. ISBN 9798400701047. doi: 10.1145/3580507.3597803. URL https://doi.org/10.1145/3580507.3597803.

Farbod Ekbatani, Rad Niazadeh, Pranav Nuti, and Jan Vondrák. Prophet inequalities with cancellation costs. In Bojan Mohar, Igor Shinkar, and Ryan O'Donnell, editors, *Proceedings of the 56th Annual ACM Symposium on Theory of Computing, STOC 2024, Vancouver, BC, Canada, June 24-28, 2024*, pages 1247–1258. ACM, 2024. doi: 10.1145/3618260.3649786. URL https://doi.org/10.1145/3618260.3649786.

Omar El Housni, Qing Feng, and Huseyin Topaloglu. Assortment optimization under the multinomial logit model with covering constraints, 2024. URL https://arxiv.org/abs/2411.10310. arXiv v2, last revised 2 Oct 2025.

Scott Fay and Jinhong Xie. Probabilistic goods: A creative way of selling products and services. *Marketing Science*, 27(4):674–690, 2008.

Scott Fay, Jinhong Xie, and Cong Feng. The effect of probabilistic selling on the optimal product mix. *Journal of Retailing*, 91(3):451–467, 2015. doi: 10.1016/j.jretai.2015.01.004.

Scott A. Fay and Jinhong Xie. Timing of product allocation: Using probabilistic selling to enhance inventory management. *Manag. Sci.*, 61:474–484, 2015. doi: 10.1287/mnsc.2014.1948.

Moran Feldman, Ola Svensson, and Rico Zenklusen. Online contention resolution schemes. In *Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '16, page 1014–1033, USA, 2016. Society for Industrial and Applied Mathematics. ISBN 9781611974331.

Yiding Feng, Rad Niazadeh, and Amin Saberi. Near-optimal bayesian online assortment of reusable resources. *Operations Research*, 72(5):1861–1873, 2024. doi: 10.1287/opre.2020.0687. URL https://doi.org/10.1287/opre.2020.0687.

Rajiv Gandhi, Samir Khuller, Srinivasan Parthasarathy, and Aravind Srinivasan. Dependent rounding and its applications to approximation algorithms. *Journal of the ACM*, 2006. doi: 10.1145/1147954.1147956. URL https://dl.acm.org/doi/10.1145/1147954.1147956.

Long Gao, Dawei Jian, Mehmet Gumus, and Birendra K. Mishra. Managing channel profits with positive demand externalities. *Management Science*, 71(10):8491–8509, 2025.

Achal Goyal and A. Sapra. Managing replenishment and clearance of perishables: Last-in, first-out (lifo) issuing policy and age-sensitive demand. *Production and Operations Management*, 33:1031 – 1052, 2024. doi: 10.1177/10591478241238970.

Peng Hu, Stephen Shum, and Man Yu. Joint inventory and markdown management for perishable goods with strategic consumer behavior. *Oper. Res.*, 64:118–134, 2015. doi: 10.1287/opre.2015.1439.

Tingliang Huang and Yimin Yu. Sell probabilistic goods? a behavioral explanation for opaque selling. *Marketing Science*, 33(5):743–759, 2014. doi: 10.1287/mksc.2014.0851. URL https://doi.org/10.1287/mksc.2014.0851.

Uzma Khan and Alexander DePaoli. Brand loyalty in the face of stockouts. *Journal of the Academy of Marketing Science*, 52(1):44–74, 2024.

Robert Kleinberg and S. Matthew Weinberg. Matroid prophet inequalities. In *Proceedings of the 44th Symposium on Theory of Computing*, pages 123–136. ACM, 2012. doi: 10.1145/2213977.2213991. URL https://doi.org/10.1145/2213977.2213991.

Qing Li, Christopher S. Tang, and He Xu. Mitigating the double-blind effect in opaque selling: Inventory and information. *Production and Operations Management*, 29:35 – 54, 2019. doi: 10.1111/poms.13087.

Vasilis Livanos and Ruta Mehta. Minimization is harder in the prophet world. In David P. Woodruff, editor, *Proceedings of the 2024 ACM-SIAM Symposium on Discrete Algorithms, SODA 2024, Alexandria, VA, USA, January 7-10, 2024*, pages 424–461. SIAM, 2024. doi: 10.1137/1.9781611977912.17. URL https://doi.org/10.1137/1.9781611977912.17.

S. Mitra. Newsvendor problem with clearance pricing. *Eur. J. Oper. Res.*, 268:193–202, 2018. doi: 10.1016/j.ejor.2018.01.023.

Steven Nahmias. Perishable inventory theory: A review. *Operations Research*, 30(4):680–708, 1982.

ReFED. How can retailers reduce food waste? https://refed.org/stakeholders/retailers/, 2024. Stakeholder recommendations for retailers. Accessed June 26, 2026.

H. Ren and Tingliang Huang. Opaque selling and inventory management in vertically differentiated markets. *Manuf. Serv. Oper. Manag.*, 24:2543–2557, 2022. doi: 10.1287/msom.2021.1041.

Claudia Rosales, Judith M. Whipple, and Jennifer Blackhurst. The impact of distribution channel decisions and repeated stockouts on manufacturer and retailer performance. *IEEE Transactions on Engineering Management*, 66(3):312–324, 2019. doi: 10.1109/TEM.2018.2835653.

E. Samuel-Cahn. Comparison of threshold stop rules and maximum for independent nonnegative random variables. *Annals of Probability*, 12:1213–1216, 1984. doi: 10.1214/aop/1176993150.

Katsunobu Sasanuma, Akira Hibiki, and Thomas Sexton. An opaque selling scheme to reduce shortage and wastage in perishable inventory systems. *Operations Research Perspectives*, 9:100220, 2022. doi: 10.1016/j.orp.2021.100220. URL https://doi.org/10.1016/j.orp.2021.100220.

Jie Sun, Yanan Wang, Chun Yang, Jiangjie Chen, Wei Wei, Wei Miao, Hanchu Sun, and Chao Gu. Is there any way to increase consumers' purchase intention regarding surplus food blind-boxes? an exploratory study. *BMC Psychology*, 12:103, 2024. doi: 10.1186/s40359-024-01587-y.

Jie Sun, Xinyi Huang, Chun Yang, Wei Wei, Wei Miao, and Chao Gu. Food waste to blind box surprises: A study on increasing consumers' purchase intention of surplus food blind boxes. *Acta Psychologica*, 261: 105787, 2025. doi: 10.1016/j.actpsy.2025.105787.

Xin Tian, Hailiang Jiang, and Xuan Zhao. Product assortment and online sales in community group-buying channel: Evidence from a convenience store chain. *Journal of Retailing and Consumer Services*, 79: 103838, 2024. doi: 10.1016/j.jretconser.2024.103838.

Too Good To Go. Smart food waste solutions for every food business. https://www.toogoodtogo.com/en-us/food-waste-solution, 2024. Accessed 25 Mar 2026.

Luyi Yang and Man Yu. Too good to go: Combating food waste with surprise clearance. *Management Science*, 2025. doi: 10.1287/mnsc.2023.03001. URL https://doi.org/10.1287/mnsc.2023.03001.

Fan Zhou, Hansheng Jiang, Andrea Li, and Joline Uichanco. Designing surprise bags for surplus foods, 2024. URL https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5002233. Available at SSRN; last revised 20 Mar 2026.

This page is intentionally blank. The e-companion (EC) starts next page.

EC.1. Further Related Work

Opaque selling and surplus-food recovery. Our work is related to probabilistic and opaque selling, where firms sell products without fully specifying their realized identities at the time of purchase [Fay and Xie, 2008, 2015; Huang and Yu, 2014]. This literature shows how delayed product assignment can pool demand uncertainty and improve inventory utilization, and later work connects opaque selling more directly to leftover inventory and perishable waste reduction [Li et al., 2019; Ren and Huang, 2022; Sasanuma et al., 2022]. Surprise-bag platforms bring this mechanism into surplus-food recovery, where retailers sell discounted opaque bags whose contents are determined from end-of-day inventory [Yang and Yu, 2025; Zhou et al., 2024]. We differ from this literature by studying the operational commitment created by these sales: bag availability must be chosen before regular demand is fully resolved, while final fulfillment must respect remaining inventory, waste costs, and bag-composition requirements.

Perishable inventory with secondary channels. Our work is also related to perishable-inventory models with markdown, clearance, donation, or salvage options. This literature studies how retailers trade off regular-price sales against spoilage, clearance revenue, and future demand cannibalization, with recovery decisions made either after leftover inventory is observed or jointly with replenishment before current demand is realized [Cachon and Kök, 2007; Goyal and Sapra, 2024; Hu et al., 2015; Mitra, 2018]. These models clarify the value of secondary outlets for distressed or aging inventory, but the secondary option is usually transparent and inventory-facing. In contrast, surprise bags create an opaque, commitment-based recovery channel: the retailer decides how many bags to offer before final surplus is known, and later assigns realized remaining products to those commitments subject to bag-composition requirements.

Online allocation and prophet inequalities. Our performance analysis is related to prophet inequalities and online allocation, where online policies are compared with a clairvoyant benchmark under uncertainty and feasibility constraints [Alaei et al., 2012; Kleinberg and Weinberg, 2012; Samuel-Cahn, 1984]. This literature develops threshold- and value-based rules for irreversible allocation decisions, with guarantees that quantify the cost of acting before future arrivals or values are known. A similar information gap drives our tight bound: the offline benchmark can route inventory after observing regular demand, whereas an online policy must commit inventory while demand is still unresolved. In our setting, however, the loss is specific to perishable surplus recovery, because the commitment trades off regular-channel upside, bag-channel recovery value, waste cost, and later fulfillment feasibility.

EC.2. Bag Requirements: Examples and Formulations

This section gives three examples of how the soft composition constraints in (1) can be specified. For clarity, we write each example as a single constraint row; multiple such requirements can be imposed simultaneously by adding rows. A generic row takes the form

$$\sum\nolimits_{i\in[N]} a_i \mathbb{E}[z_i] \le b, \tag{EC.1}$$

where $a_i \in \mathbb{R}$ and $b \ge 0$. The examples below show how average-quality floor, product-availability caps, and category-concentration limits requirements fit this form.

Average-quality floor. Surprise-bag quality can affect consumers' purchase and repurchase intentions, and realized bag experiences can shape future demand through customer ratings [Sun et al., 2024, 2025; Zhou et al., 2024]. Let u_i denote a quality score for product i. The retailer may require the ratio of expected total quality to the expected total number of units allocated to surprise bags to be at least $\overline{u}$. When this expected total is positive, the requirement is

$$\frac{\sum_{i\in[N]} u_i \mathbb{E}[z_i]}{\sum_{i\in[N]} \mathbb{E}[z_i]} \ge \overline{u}, \qquad \text{or equivalently} \qquad \sum\nolimits_{i\in[N]} (\overline{u} - u_i)\mathbb{E}[z_i] \le 0.$$

The linear form remains well defined and is satisfied when no surprise-bag orders are accepted. It therefore fits (EC.1) with $a_i = \overline{u} - u_i$ and $b = 0$.

Product-availability caps. Using too much of a particular product for surprise-bag fulfillment may reduce its availability in the regular channel, where stockouts can affect customer loyalty and future demand [Gao et al., 2025; Khan and DePaoli, 2024; Rosales et al., 2019]. To protect availability for regular sales, the retailer may cap the expected number of units of product $i \in [N]$ used for surprise-bag fulfillment at some limit $z_i^{\max} \ge 0$. The requirement is

$$\mathbb{E}[z_i] \le z_i^{\max}.$$

This fits (EC.1) with $a_{i'} = \mathbb{1}\{i' = i\}$ for $i' \in [N]$ and $b = z_i^{\max}$.

Category-concentration limits. An overly concentrated surprise-bag mix may reduce product variety and make the offering less attractive to consumers. Evidence from surplus-food blind-box settings suggests that greater product-type variety can increase consumer interest [Sun et al., 2024], while related retail research links assortment diversity to stronger sales performance [Tian et al., 2024]. To maintain a diversified bag mix, the retailer may require the ratio of the expected number of units from a category $\mathcal{C} \subseteq [N]$ to the expected total number of units allocated to surprise bags to be at most $\gamma \in (0, 1)$.

When this expected total is positive, this requirement can be written as

$$\frac{\sum_{i\in\mathcal{C}} \mathbb{E}[z_i]}{\sum_{i\in[N]} \mathbb{E}[z_i]} \le \gamma, \qquad \text{or equivalently} \qquad (1-\gamma)\sum\nolimits_{i\in\mathcal{C}} \mathbb{E}[z_i] - \gamma \sum\nolimits_{i\notin\mathcal{C}} \mathbb{E}[z_i] \le 0.$$

When this expected total is zero, the linear form remains well defined and is automatically satisfied. This fits (EC.1) with

$$a_i = (1-\gamma)\mathbb{1}\{i \in \mathcal{C}\} - \gamma\mathbb{1}\{i \notin \mathcal{C}\}, \qquad b = 0.$$

EC.3. Negative Soft-Budget Limits and Competitive-Ratio Collapse

Requiring $b_\ell \geq 0$ for every ℓ in (1) ensures that accepting no surprise-bag orders satisfies every soft composition constraint. Allowing a negative soft budget can yield a soft composition constraint equivalent to requiring a minimum expected number of units to be allocated to the surprise-bag channel. The following result shows that this possibility rules out any positive competitive guarantee.

PROPOSITION EC.1. *For every $\alpha > 0$, there exists an instance I satisfying the productwise regular-channel profitability assumption (Assumption 1) and having one negative soft budget such that every feasible online policy Π satisfies*

$$\mathbb{E}[\textsc{Profit}[\Pi; I]] < \alpha\,\mathbb{E}[\textsc{Profit}[\textsc{Opt}; I]]\,.$$

Proof. Fix $\alpha > 0$ and choose $\delta \in (0, \min\{1, \alpha/2\})$. Consider an instance with one product, one period, and one unit of inventory. Let the regular reward be $r = 1/\delta^2$, the surprise-bag reward be $s = 1$, and the waste cost be $w = 0$, and let regular demand equal one with probability δ and zero otherwise. Set both the regular-display cap and the bag-offer cap to one. Impose the soft composition constraint

$$-\mathbb{E}[z_1] \leq -(1-\delta),$$

whose soft budget is negative. Equivalently, every feasible policy must satisfy $\mathbb{E}[z_1] \geq 1-\delta$.

Consider first any feasible online policy. Its surprise-bag decision is made before regular demand is observed. Since an accepted surprise-bag order must be fulfilled, the single unit cannot simultaneously be committed to a bag and exposed to regular demand. The soft constraint therefore requires the policy to make a surprise bag available with probability at least $1-\delta$. To ensure fulfillment, making a surprise bag available requires withholding the unit from the regular channel. The unit is therefore available to the regular channel with probability at most δ. Since this decision is made before demand is realized, the probability that the unit is both available to the regular channel and regular demand equals one is at most δ^2. Hence, its expected surprise-bag revenue is at most 1, while its expected regular revenue is also at most $\delta^2(1/\delta^2) = 1$. Thus,

$$\mathbb{E}[\textsc{Profit}[\Pi; I]] \leq 2.$$

The offline optimal policy observes demand before making its decisions. It uses the unit for surprise-bag fulfillment when demand is zero and sells the unit through the regular channel when demand is one. Its

expected number of units allocated to surprise bags is therefore $1-\delta$, so the soft constraint is satisfied with equality. Its expected profit is

$$\mathbb{E}[\textsc{Profit}[\textsc{Opt}; I]] = 1 - \delta + \frac{1}{\delta} > \frac{2}{\alpha}.$$

Finally, the regular-only greedy policy earns expected profit $\delta r = 1/\delta > 0$, so the instance satisfies Assumption 1. This proves the claim.

□

The loss of any positive competitive guarantee is driven by the interaction between a negative soft budget and the information structure. A negative soft budget can turn a composition restriction into a requirement that a minimum expected number of units be allocated to surprise bags, forcing an online retailer to commit inventory to surprise-bag fulfillment before regular demand is known. The offline optimal policy can instead satisfy this requirement selectively in low-demand realizations while preserving inventory for rare, high-value regular sales. The resulting information advantage can be made arbitrarily large.

EC.4. Extension to m-Unit Surprise Bags

We extend the LP-based reservation policy to surprise bags that contain exactly m physical units instead of a single unit. We retain the aggregate composition model of Section 2, which applies only to the expected product allocation across all bags and therefore creates no within-bag packing problem. Under this aggregate-composition model, the additional feasibility issue is that every realized reservation must contain a number of units divisible by m and therefore correspond to an integral number of complete bags. We formulate a planning problem, implement its solution without loss, and establish the optimal competitive ratio $1/(2+\kappa_m)$, where $\kappa_m \triangleq mw/s$.

Except that each bag contains exactly m physical units, the model of Section 2 is unchanged. In particular, its primitives, timing and information structure, regular-sales dynamics, and offline benchmark continue to apply. The soft aggregate composition constraints remain

$$\sum\nolimits_{i\in[N]} a_{i,\ell}\mathbb{E}[z_i] \le b_\ell, \qquad \ell \in [L].$$

In the one-unit static planning problem, a single reservation vector suffices: each reserved unit supports one surprise bag, so the total number of reserved units also determines the total number of bags that can be made available. With m-unit bags, this simplification no longer applies. If the plan makes h bags available, it must reserve exactly mh physical units. The extension therefore separates the bag count from the unit-level reservation, as enforced by the branch-cardinality constraints below. At the policy level, the pathwise fulfillment identity is therefore

$$\sum\nolimits_{i\in[N]} z_i = m \sum\nolimits_{t\in[T]} y_t.$$

Both the online policy and the offline benchmark are subject to the displayed fulfillment identity for m-unit bags. The profitability assumption (Assumption 1) and competitive-ratio definition (Definition 1) continue to apply. Because s is now earned jointly by m units, we compare the regular reward with the per-unit bag reward and assume

$$r_i \geq \frac{s}{m} \geq w, \qquad i \in [N], \tag{EC.2}$$

which replaces the one-unit ordering $r_i > s \geq w$ in Section 2.

EC.4.1. Planning the Reservation

We continue to use the regular-only baseline, unit indexing, regular values $v_{i,k}$, and unit set $\mathcal{U}$ from Section 3.1 to formulate the reservation plan. The regular-only baseline satisfies

$$\mathbb{E}[\textsc{Profit}[\text{RO}]] = \sum\nolimits_{(i,k)\in\mathcal{U}} v_{i,k}.$$

A natural first attempt is to retain the one-unit reservation LP, value each reserved unit at s/m, and increase its capacity to mTY, since each bag requires m selected units. Such a formulation controls only the expected number of reserved units. After dependent rounding, however, the realized number of reserved units need not be divisible by m, so the reservation may not correspond to an integral number of bags.

We resolve this issue by conditioning the reservation on an integral total bag count. The total number of surprise bags made available over the horizon can be at most $h^{\max} \triangleq \min\left\{TY, \left\lfloor \sum_{i\in[N]} q_i/m \right\rfloor\right\}$. This bound follows because the per-period cap permits at most TY bags, and the inventory can supply at most $\left\lfloor \sum_{i\in[N]} q_i/m \right\rfloor$ complete bags. For each $h \in \{0,\dots,h^{\max}\}$, let branch h represent making h surprise bags available over the horizon, and let ρ_h be the probability that the reservation plan selects this branch.

Conditional on selecting branch h, the formulation assigns a fractional reservation with total mass mh. For each $(i,k) \in \mathcal{U}$, let $u^h_{i,k}$ be the joint probability that the plan selects h and reserves unit (i,k). Thus, $\sum_{h=0}^{h^{\max}} u^h_{i,k}$ is the marginal probability that unit (i,k) is reserved. This gives the following static linear program:

$$\begin{aligned}
\mathbb{E}[\textsc{Profit}[\text{RO}]] + \max_{\rho,u} \quad & \sum\nolimits_{h=0}^{h^{\max}} \left(sh\rho_h - \sum\nolimits_{(i,k)\in\mathcal{U}} v_{i,k} u^h_{i,k} \right) \\
\text{s.t.} \quad & \sum\nolimits_{h=0}^{h^{\max}} \rho_h = 1, \\
& \rho_h \geq 0, \qquad h = 0,\dots,h^{\max}, \\
& \sum\nolimits_{(i,k)\in\mathcal{U}} u^h_{i,k} = mh\rho_h, \qquad h = 0,\dots,h^{\max}, \\
& 0 \leq u^h_{i,k} \leq \rho_h, \qquad (i,k)\in\mathcal{U},\ h = 0,\dots,h^{\max}, \\
& \sum\nolimits_{h=0}^{h^{\max}} \sum\nolimits_{(i,k)\in\mathcal{U}} a_{i,\ell} u^h_{i,k} \leq b_\ell, \qquad \ell \in [L],
\end{aligned} \tag{LP$^m_{\text{st}}$}$$

The first two constraints specify a randomized integral bag count. The third ensures that, conditional on branch h, the fractional reservation has total mass mh. The fourth makes the joint reservation probabilities valid, and the fifth retains the aggregate composition requirements across branches.

To describe the fractional reservation conditional on selecting branch h, for each branch h with $\rho_h > 0$, define $\overline{u}^h$ coordinatewise by

$$\overline{u}^h_{i,k} \triangleq \frac{u^h_{i,k}}{\rho_h}, \qquad (i,k) \in \mathcal{U}. \tag{EC.3}$$

The following lemma shows that an optimal solution can be chosen so that these conditional reservations are suffix-form.

LEMMA EC.1 **(Branchwise suffix-form optimal solution).** *There is an optimal solution to* ($\mathrm{LP}^m_{\mathrm{st}}$)*, denoted* $(\{\rho_h\}_{h=0}^{h^{\max}}, \{u^h_{i,k}\}_{h=0,\dots,h^{\max},(i,k)\in\mathcal{U}})$*, such that, for every branch* h *with* $\rho_h > 0$ *and every product* i*, there exist* $k^h_i \in [q_i]$ *and* $\theta^h_i \in [0,1]$ *satisfying*

$$\overline{u}^h_{i,k} = \begin{cases} 0, & k < k^h_i, \\ \theta^h_i, & k = k^h_i, \\ 1, & k > k^h_i, \end{cases}$$

where $\overline{u}^h_{i,k}$ *is defined in Equation* (EC.3).

Proof. Fix an optimal solution. For any branch h with $\rho_h > 0$, any product i, and indices $k < k'$ such that $u^h_{i,k} > 0$ and $u^h_{i,k'} < \rho_h$, choose the smallest such k and the largest such k'. Move $\epsilon = \min\{u^h_{i,k},\ \rho_h - u^h_{i,k'}\}$ units of reservation mass from (i,k) to (i,k'). All units of product i enter the constraints with the same coefficients, so feasibility is preserved. Since $-v_{i,k}$ is nondecreasing in k, the objective does not decrease. Each exchange either sets the chosen lower-index coordinate to zero or the chosen higher-index coordinate to ρ_h. Hence, after finitely many exchanges, no such pair remains. The coordinates of $\overline{u}^h$ for each product then form a suffix, with at most one fractional coordinate, which gives the stated form. □

EC.4.2. Implementing the Reservation Plan

We apply the dependent-rounding procedure of Chekuri et al. [2010]; Gandhi et al. [2006], described in Section 3.2. The new step is to first sample a bag-count branch and then round $\overline{u}^h$. As in Section 3.2, dependent rounding updates only fractional coordinates and leaves every integral coordinate unchanged. Moreover, because $\overline{u}^h$ has integral total mass mh, the exact-cardinality property in Lemma 2 ensures that the rounded reservation contains exactly mh units almost surely. The complete policy is given in Algorithm 4.

LEMMA EC.2. *The* m*-unit LP-based reservation policy, denoted* SR^m*, is feasible and has expected profit equal to the optimal value of* ($\mathrm{LP}^m_{\mathrm{st}}$)*, i.e.,*

$$\mathbb{E}[\textsc{Profit}[\mathrm{SR}^m]] = \mathrm{LP}^m_{\mathrm{st}}.$$

Algorithm 4: m-Unit LP-Based Reservation (SR^m) Policy

Find a suffix-form optimal solution $(\{\rho_h\}_{h=0}^{h^{\max}}, \{u_{i,k}^h\}_{h=0,\dots,h^{\max},(i,k)\in\mathcal{U}})$ of ($\mathrm{LP}_{\mathrm{st}}^m$).

Sample h according to $\{\rho_h\}_{h=0}^{h^{\max}}$ and set $(\overline{u}^h)' \leftarrow \textsc{DepRound}(\overline{u}^h)$, where $\overline{u}^h$ is defined in (EC.3).

Set $(\overline{u}^h)'_i \leftarrow \sum_{k\in[q_i]} (\overline{u}^h)'_{i,k}$ for each $i \in [N]$.

Initialize $q_{i,0}^R \leftarrow q_i - (\overline{u}^h)'_i$ for each $i \in [N]$, and $q_0^S \leftarrow h$.

for $t = 1, \dots, T$ **do**

  Set $x_{i,t} \leftarrow \min\{X_i, q_{i,t-1}^R\}$ for each $i \in [N]$, and $y_t \leftarrow \min\{Y, q_{t-1}^S\}$.

  `/* Regular demand in period t is realized. */`

  Update $q_{i,t}^R \leftarrow q_{i,t-1}^R - \min\{d_{i,t}, x_{i,t}\}$ for each $i \in [N]$, and $q_t^S \leftarrow q_{t-1}^S - y_t$.

Partition the mh selected units arbitrarily into h bags of exactly m units, and use these bags to fulfill the accepted surprise-bag orders.

Proof. We first show that the policy is feasible. The rounding is performed before the horizon and independently of demand. Conditional on the sampled branch h, dependent rounding selects exactly mh units. The policy respects the display-cap and bag-offer constraints. Since $h \le h^{\max} \le TY$, the bag-offer rule in Algorithm 4 makes all h bags available by the end of the horizon. The selected units are never displayed in the regular channel and therefore remain available for fulfillment; partitioning them arbitrarily into h groups of size m fulfills every accepted order.

To verify the soft composition constraints, conditional on branch h, for every product $i \in [N]$,

$$z_i = \sum\nolimits_{k\in[q_i]} (\overline{u}^h)'_{i,k}.$$

By marginal preservation in Lemma 2, $\mathbb{E}\big[(\overline{u}^h)'_{i,k} \mid h\big] = \overline{u}_{i,k}^h$ for every $(i,k) \in \mathcal{U}$. Thus,

$$\mathbb{E}[z_i] = \sum\nolimits_{h=0}^{h^{\max}} \sum\nolimits_{k\in[q_i]} \rho_h \mathbb{E}\big[(\overline{u}^h)'_{i,k} \mid h\big] = \sum\nolimits_{h=0}^{h^{\max}} \sum\nolimits_{k\in[q_i]} u_{i,k}^h.$$

Feasibility of (ρ, u) in ($\mathrm{LP}_{\mathrm{st}}^m$) then implies that the soft composition constraints are satisfied.

We next consider profit. For each product, only its fractional boundary coordinate can change under dependent rounding. Hence every realized reservation remains suffix-form. Conditional on a realized $(h, (\overline{u}^h)')$, the selected units correspond to h complete bags and therefore earn sh. The unreserved units form productwise prefixes and contribute their corresponding regular values. Hence

$$\mathbb{E}\big[\textsc{Profit}[\mathrm{SR}^m] \mid h, (\overline{u}^h)'\big] = \mathbb{E}[\textsc{Profit}[\mathrm{RO}]] + sh - \sum\nolimits_{(i,k)\in\mathcal{U}} v_{i,k} (\overline{u}^h)'_{i,k}.$$

Taking expectations over the rounding and the sampled branch and applying $\rho_h \mathbb{E}\big[(\overline{u}^h)'_{i,k} \mid h\big] = u_{i,k}^h$ again yields $\mathrm{LP}_{\mathrm{st}}^m$. □

EC.4.3. Competitive Performance

The following result extends the one-unit guarantee in Theorem 1 to m-unit bags.

THEOREM EC.1 **(Optimal competitive ratio for m-unit bags).** *For every m, let $\kappa_m \triangleq mw/s \in [0,1]$. Under Assumption 1 and (EC.2), the m-unit LP-based reservation policy achieves competitive ratio $\frac{1}{2+\kappa_m}$. In particular, it is $1/3$-competitive over all instances satisfying the two assumptions.*

Moreover, this guarantee is optimal for every $\kappa_m \in [0,1]$: for every $\varepsilon > 0$, there exists a m-unit instance satisfying Assumption 1 and (EC.2) with $mw/s = \kappa_m$ such that every feasible online policy earns at most $\frac{1}{2+\kappa_m} + \varepsilon$ times the offline benchmark.

Proof. We first prove that the m-unit LP-based reservation policy achieves competitive ratio $1/(2+\kappa_m)$, and then establish the matching lower bound. The competitive-ratio proof has three steps. We first derive an offline upper bound using the multi-unit uniform-gain LP, denoted by $\mathrm{LP}^m_{\mathrm{ug}}$. We then compare this relaxation with the multi-unit static linear program $\mathrm{LP}^m_{\mathrm{st}}$ through a common conditional-suffix-form solution. Finally, we combine the two bounds with the lossless implementation in Lemma EC.2.

Offline upper bound. As in the one-unit setting, we upper-bound the offline benchmark by replacing the reservation value of each bag with its maximum possible incremental gain over regular-only profit. A bag earns s and avoids the waste cost on its m units. Accordingly, each bag contributes at most $s+mw$ to the improvement over the regular-only baseline. The resulting multi-unit uniform-gain LP is

$$
\begin{aligned}
\mathbb{E}[\textsc{Profit}[\mathrm{RO}]] + \max_{\rho,u} \quad & (s+mw)\sum\nolimits_{h=0}^{h^{\max}} h\rho_h \\
\text{s.t.} \quad & \sum\nolimits_{h=0}^{h^{\max}} \rho_h = 1, \\
& \rho_h \ge 0, \qquad h = 0,\dots,h^{\max}, \\
& \sum\nolimits_{(i,k)\in\mathcal{U}} u^h_{i,k} = mh\rho_h, \qquad h = 0,\dots,h^{\max}, \\
& 0 \le u^h_{i,k} \le \rho_h, \qquad (i,k)\in\mathcal{U},\ h = 0,\dots,h^{\max}, \\
& \sum\nolimits_{h=0}^{h^{\max}} \sum\nolimits_{(i,k)\in\mathcal{U}} a_{i,\ell} u^h_{i,k} \le b_\ell, \qquad \ell \in [L].
\end{aligned}
\tag{$\mathrm{LP}^m_{\mathrm{ug}}$}
$$

Fix a demand realization. Let χ_i^{OFF} and χ_i be the regular sales of product i under the offline optimal policy and regular-only greedy policy, respectively, and let h^{OFF} be the number of bags made available, and hence sold, by the offline optimal policy. The regular-only policy displays every available unit up to the display cap. Hence, for each product i,

$$
\chi_i^{\mathrm{OFF}} \le \min\left\{ q_i, \sum\nolimits_{t\in[T]} \min\{d_{i,t}, X_i\} \right\} = \chi_i.
$$

Since the offline optimal policy allocates exactly mh^{OFF} units to bags,

$$
\begin{aligned}
\textsc{Profit}[\mathrm{OPT}] - \textsc{Profit}[\mathrm{RO}] &= \sum\nolimits_{i\in[N]} (r_i + w)(\chi_i^{\mathrm{OFF}} - \chi_i) + (s+mw)h^{\mathrm{OFF}} \\
&\le (s+mw)h^{\mathrm{OFF}}.
\end{aligned}
$$

On each realization, label the offline bag units of product i as the last z_i^{OFF} unit indices of that product. Set $\rho_h = \Pr[h^{\mathrm{OFF}} = h]$ and let $u_{i,k}^h$ be the joint probability that $h^{\mathrm{OFF}} = h$ and the labeled suffix contains (i,k). These variables satisfy every constraint in ($\mathrm{LP}_{\mathrm{ug}}^m$), including the aggregate composition rows. Taking expectations in the pathwise inequality and evaluating ($\mathrm{LP}_{\mathrm{ug}}^m$) at these feasible variables gives

$$\mathbb{E}[\textsc{Profit}[\textsc{Opt}]] \le \mathrm{LP}_{\mathrm{ug}}^m. \tag{EC.4}$$

LP comparison. By (EC.4) and the lossless implementation in Lemma EC.2, it suffices to establish

$$\mathrm{LP}_{\mathrm{st}}^m \ge \max\left\{\mathbb{E}[\textsc{Profit}[\mathrm{RO}]],\ \frac{1}{1+\kappa_m}\left(\mathrm{LP}_{\mathrm{ug}}^m - \mathbb{E}[\textsc{Profit}[\mathrm{RO}]]\right)\right\}. \tag{EC.5}$$

The constraints $\sum_{(i,k)\in\mathcal{U}} u_{i,k}^h = mh\rho_h$ for $h = 0,\dots,h^{\max}$ in the multi-unit static linear program couple the bag-count probabilities and unit-reservation variables. We therefore compare the two primal relaxations directly through a common conditional-suffix-form solution. Take an optimal solution (ρ^*, u^*) to ($\mathrm{LP}_{\mathrm{ug}}^m$); write its components as ρ_h^* and $(u_{i,k}^h)^*$ for $h = 0,\dots,h^{\max}$ and $(i,k)\in\mathcal{U}$. For every h with $\rho_h^* > 0$, define

$$(\overline{u}_{i,k}^h)^* \triangleq \frac{(u_{i,k}^h)^*}{\rho_h^*}, \qquad (i,k)\in\mathcal{U}.$$

By the same exchange argument as in Lemma EC.1, this solution may be chosen so that these conditional vectors are suffix-form. The solution (ρ^*, u^*) is feasible for ($\mathrm{LP}_{\mathrm{st}}^m$). To lower-bound its value in that linear program, we first show that

$$\sum_{h=0}^{h^{\max}} \sum_{(i,k)\in\mathcal{U}} v_{i,k}(u_{i,k}^h)^* \le \mathbb{E}[\textsc{Profit}[\mathrm{RO}]]. \tag{EC.6}$$

By construction, for every product i, $\{v_{i,k}\}_{k=1}^{q_i}$ is nonincreasing. Moreover, Assumption 1 implies $\sum_{k=1}^{q_i} v_{i,k} \ge 0$. For each $j\in[q_i]$,

$$\sum_{k=1}^{j} v_{i,k} \ge 0;$$

otherwise, $v_{i,j} < 0$, so all later values would also be negative, contradicting $\sum_{k=1}^{q_i} v_{i,k} \ge 0$. Now fix h with $\rho_h^* > 0$. The conditional vector defined above is suffix-form, so its complement equals one on a set of lower-indexed units, zero on all higher-indexed units, and may be fractional for one intermediate unit. Its value is therefore a convex combination of the values of the two integer prefixes obtained by setting that coordinate to zero or one. Hence

$$\sum_{k=1}^{q_i} v_{i,k}\left(1 - (\overline{u}_{i,k}^h)^*\right) \ge 0.$$

Equivalently,

$$\sum_{k=1}^{q_i} v_{i,k}(\overline{u}_{i,k}^h)^* \le \sum_{k=1}^{q_i} v_{i,k}.$$

Multiplying by ρ_h^*, summing over products and branches, and using $\sum_{h=0}^{h^{\max}} \rho_h^* = 1$ proves (EC.6).

Then, the objective of ($\mathrm{LP}^m_{\mathrm{st}}$), evaluated at (ρ^*, u^*), is at least

$$\begin{aligned}
&\mathbb{E}[\textsc{Profit}[\mathrm{RO}]] + s\sum\nolimits_{h=0}^{h^{\max}} h\rho^*_h - \sum\nolimits_{h=0}^{h^{\max}} \sum\nolimits_{(i,k)\in\mathcal{U}} v_{i,k}(u^h_{i,k})^* \\
&\geq s\sum\nolimits_{h=0}^{h^{\max}} h\rho^*_h \\
&= \frac{s}{s+mw}\left(\mathrm{LP}^m_{\mathrm{ug}} - \mathbb{E}[\textsc{Profit}[\mathrm{RO}]]\right),
\end{aligned}$$

where the inequality follows from (EC.6), and the equality follows from the objective of ($\mathrm{LP}^m_{\mathrm{ug}}$) at (ρ^*, u^*). By the definition of κ_m, $s/(s+mw) = 1/(1+\kappa_m)$. Since (ρ^*, u^*) is feasible for ($\mathrm{LP}^m_{\mathrm{st}}$),

$$\mathrm{LP}^m_{\mathrm{st}} \geq \frac{1}{1+\kappa_m}\left(\mathrm{LP}^m_{\mathrm{ug}} - \mathbb{E}[\textsc{Profit}[\mathrm{RO}]]\right).$$

Because $b_\ell \geq 0$ for every composition row, the zero-bag solution is feasible. Hence,

$$\mathrm{LP}^m_{\mathrm{st}} \geq \mathbb{E}[\textsc{Profit}[\mathrm{RO}]].$$

Combining these two lower bounds on $\mathrm{LP}^m_{\mathrm{st}}$ and using the lossless implementation in Lemma EC.2 gives

$$\mathbb{E}[\textsc{Profit}[\mathrm{SR}^m]] \geq \max\left\{\mathbb{E}[\textsc{Profit}[\mathrm{RO}]],\ \frac{1}{1+\kappa_m}\left(\mathrm{LP}^m_{\mathrm{ug}} - \mathbb{E}[\textsc{Profit}[\mathrm{RO}]]\right)\right\}. \tag{EC.7}$$

Putting the bounds together. By Assumption 1 and feasibility of the zero-bag solution in ($\mathrm{LP}^m_{\mathrm{ug}}$), respectively,

$$\mathbb{E}[\textsc{Profit}[\mathrm{RO}]] \geq 0, \qquad \mathrm{LP}^m_{\mathrm{ug}} - \mathbb{E}[\textsc{Profit}[\mathrm{RO}]] \geq 0.$$

For any $a, b, \kappa \geq 0$,

$$a + b \leq (2+\kappa)\max\left\{a, \frac{b}{1+\kappa}\right\}.$$

Applying this inequality with $a = \mathbb{E}[\textsc{Profit}[\mathrm{RO}]]$, $b = \mathrm{LP}^m_{\mathrm{ug}} - \mathbb{E}[\textsc{Profit}[\mathrm{RO}]]$, and $\kappa = \kappa_m$ gives

$$\begin{aligned}
&\mathbb{E}[\textsc{Profit}[\mathrm{RO}]] + \left(\mathrm{LP}^m_{\mathrm{ug}} - \mathbb{E}[\textsc{Profit}[\mathrm{RO}]]\right) \\
&\leq (2+\kappa_m)\max\left\{\mathbb{E}[\textsc{Profit}[\mathrm{RO}]],\ \frac{1}{1+\kappa_m}\left(\mathrm{LP}^m_{\mathrm{ug}} - \mathbb{E}[\textsc{Profit}[\mathrm{RO}]]\right)\right\}.
\end{aligned} \tag{EC.8}$$

Combining (EC.4), (EC.7), and (EC.8) yields

$$\begin{aligned}
\mathbb{E}[\textsc{Profit}[\mathrm{Opt}]] &\leq \mathrm{LP}^m_{\mathrm{ug}} \\
&= \mathbb{E}[\textsc{Profit}[\mathrm{RO}]] + \left(\mathrm{LP}^m_{\mathrm{ug}} - \mathbb{E}[\textsc{Profit}[\mathrm{RO}]]\right) \\
&\leq (2+\kappa_m)\max\left\{\mathbb{E}[\textsc{Profit}[\mathrm{RO}]],\ \frac{1}{1+\kappa_m}\left(\mathrm{LP}^m_{\mathrm{ug}} - \mathbb{E}[\textsc{Profit}[\mathrm{RO}]]\right)\right\} \\
&\leq (2+\kappa_m)\mathbb{E}[\textsc{Profit}[\mathrm{SR}^m]],
\end{aligned}$$

which proves the guarantee.

For the matching impossibility result, consider an instance with one product, one period, m units, no composition constraints, and $X_1 = m$ and $Y = 1$. Let

$$s = 1, \qquad w = \frac{\kappa_m}{m}, \qquad \delta = \frac{1+\kappa_m}{m(r+w)}$$

where $r > 1$, and let regular demand equal m with probability δ and zero otherwise. If a surprise bag is made available, it must be fulfilled. Hence, a feasible online policy cannot make a bag available and expose any of the m units to regular demand simultaneously. Leaving some units undisplayed without making a bag available is dominated by displaying them, since each additional regular display has expected incremental profit $\delta(r+w) > 0$. The only undominated online decisions are therefore to make a bag available or leave all m units in the regular channel. Their expected profits are, respectively,

$$1 \qquad \text{and} \qquad \delta mr - (1-\delta)mw = 1.$$

Thus, every randomized online policy earns expected profit at most 1. The offline optimal policy observes regular demand before making this decision: it chooses the regular channel when demand is m and the surprise-bag channel otherwise, so the offline benchmark is

$$\delta mr + (1-\delta) \longrightarrow 2 + \kappa_m \qquad \text{as } r \to \infty.$$

Hence the competitive ratio approaches $1/(2+\kappa_m)$. The construction satisfies Assumption 1 because the regular-only greedy policy has expected profit 1, and (EC.2) holds for sufficiently large r. Choosing r sufficiently large proves the matching lower bound. □

Relation to the one-unit model. When $m = 1$, $\kappa_m = w/s$, so we recover the guarantee and matching lower bound of Theorem 1.

Scope of the extension. This extension imposes composition requirements only on the aggregate expected allocation. Requirements on the contents of each individual bag would couple reservation with a separate combinatorial packing problem and are not covered by Theorem EC.1.

EC.5. Proofs for Section 3

EC.5.1. Proof of Lemma 2

Proof. Consider one pairwise update applied to fractional coordinates i and j. Conditional on the current vector y, the expected value of coordinate i after the update is

$$\frac{\delta^-}{\delta^+ + \delta^-}(y_i + \delta^+) + \frac{\delta^+}{\delta^+ + \delta^-}(y_i - \delta^-) = y_i.$$

The same calculation applies to coordinate j, while all other coordinates remain unchanged. Thus, every pairwise update preserves each coordinate in conditional expectation. If a final Bernoulli rounding is needed,

it also preserves the remaining fractional coordinate in conditional expectation. Repeated application of the law of iterated expectations therefore gives $\mathbb{E}[y_j] = x_j$ for every $j \in J$.

Each pairwise update also preserves $y_i + y_j$ exactly and, by the definitions of δ^+ and δ^-, makes at least one of the two selected coordinates integral. Hence, the procedure terminates with at most one fractional coordinate, while the total remains $\sum_{j\in J} x_j$.

If no fractional coordinate remains, the result follows. Otherwise, let i be the remaining fractional coordinate. Since all other coordinates are integral,

$$y_i = \sum\nolimits_{j\in J} x_j - \left\lfloor \sum\nolimits_{j\in J} x_j \right\rfloor .$$

The final Bernoulli draw therefore rounds the total to either $\lfloor \sum_{j\in J} x_j \rfloor$ or $\lceil \sum_{j\in J} x_j \rceil$, completing the proof. □

EC.5.2. Proof of the Competitive Guarantee in Theorem 1

We first upper-bound the offline benchmark by the following uniform-gain LP:

$$\begin{aligned} \mathbb{E}[\textsc{Profit}[\mathrm{RO}; I]] + \max_{u\in[0,1]^{\mathcal{U}}} & \sum\nolimits_{(i,k)\in\mathcal{U}} (s+w) u_{i,k} \\ \text{s.t.} \quad & \sum\nolimits_{(i,k)\in\mathcal{U}} u_{i,k} \le TY, \\ & \sum\nolimits_{(i,k)\in\mathcal{U}} a_{i,\ell} u_{i,k} \le b_\ell, \qquad \ell \in [L]. \end{aligned} \tag{$\mathrm{LP}_{\mathrm{ug}}$}$$

Let $z_i^\star$ be the number of units of product i allocated to surprise bags by $\textsc{Opt}$. For each product i, choose a suffix-form vector $\{u_{i,k}\}_{k\in[q_i]} \in [0,1]^{q_i}$ such that

$$\sum\nolimits_{k\in[q_i]} u_{i,k} = \mathbb{E}[z_i^\star] .$$

Such a vector exists because $0 \le \mathbb{E}[z_i^\star] \le q_i$. By feasibility of the offline optimal policy, this vector satisfies the bag-capacity and composition constraints of $\mathrm{LP}_{\mathrm{ug}}$.

On every demand realization, the regular-only greedy policy makes at least as many regular sales of each product as the offline optimal policy. Relative to this baseline, assigning a unit to a surprise bag yields at most $s+w$: this gain occurs when the unit would otherwise remain unsold, whereas assigning a unit that would sell regularly changes its contribution by $s - r_i < 0$. Hence,

$$\textsc{Profit}[\textsc{Opt}; I] \le \textsc{Profit}[\mathrm{RO}; I] + (s+w) \sum\nolimits_{i\in[N]} z_i^\star .$$

Taking expectations, the right-hand side equals the objective of $\mathrm{LP}_{\mathrm{ug}}$ at the feasible vector above. Therefore,

$$\mathrm{LP}_{\mathrm{ug}} \ge \mathbb{E}[\textsc{Profit}[\textsc{Opt}; I]] . \tag{EC.9}$$

Proof of Theorem 1. By (EC.9), it remains to prove

$$(2+\kappa)\mathrm{LP}_{\mathrm{st}} \geq \mathrm{LP}_{\mathrm{ug}}.$$

To prove this, we compare the dual forms of $\mathrm{LP}_{\mathrm{st}}$ and $\mathrm{LP}_{\mathrm{ug}}$ on a common augmented row set, with row 0 for the bag-capacity constraint and rows $1,\dots,L$ for the soft-constraint rows. For any product $i \in [N]$, let $\mathbf{a}_i \in \mathbb{R}^{L+1}$ denote the common coefficient vector of its units in this augmented system, with $(\mathbf{a}_i)_0 = 1$ and $(\mathbf{a}_i)_\ell = a_{i,\ell}$ for $\ell \in [L]$. Let $\mathbf{b}$ be the corresponding right-hand-side vector, with $(\mathbf{b})_0 = TY \geq 0$ and $(\mathbf{b})_\ell = b_\ell \geq 0$ for $\ell \in [L]$. Using the dual forms derived in Section EC.8, we have

$$\mathrm{LP}_{\mathrm{st}} = \mathbb{E}[\textsc{Profit}[\mathrm{RO}; I]] + \min_{\lambda \in [0,\infty)^{L+1}} \left\{ \mathbf{b}^\top \lambda + \sum\nolimits_{(i,k)\in\mathcal{U}} \left(s - v_{i,k} - \mathbf{a}_i^\top \lambda\right)^+ \right\},$$
$$\mathrm{LP}_{\mathrm{ug}} = \mathbb{E}[\textsc{Profit}[\mathrm{RO}; I]] + \min_{\mu \in [0,\infty)^{L+1}} \left\{ \mathbf{b}^\top \mu + \sum\nolimits_{(i,k)\in\mathcal{U}} \left(s + w - \mathbf{a}_i^\top \mu\right)^+ \right\}.$$

Let λ^* be an optimal dual solution for $\mathrm{LP}_{\mathrm{st}}$, and set $\mu \triangleq (1+\kappa)\lambda^*$. Since $\lambda^* \geq 0$, μ is feasible for the dual of $\mathrm{LP}_{\mathrm{ug}}$. Therefore, evaluating the dual objective of $\mathrm{LP}_{\mathrm{ug}}$ at μ gives

$$\mathrm{LP}_{\mathrm{ug}} \leq \mathbb{E}[\textsc{Profit}[\mathrm{RO}; I]] + \mathbf{b}^\top \mu + \sum\nolimits_{(i,k)\in\mathcal{U}} \left(s + w - \mathbf{a}_i^\top \mu\right)^+. \tag{EC.10}$$

Moreover, since $(\mathbf{b})_\ell \geq 0$ for all $\ell \in \{0\} \cup [L]$,

$$\mathbf{b}^\top \mu = (1+\kappa)\mathbf{b}^\top \lambda^* \leq (2+\kappa)\mathbf{b}^\top \lambda^*. \tag{EC.11}$$

For any product $i \in [N]$, denote $\alpha_i \triangleq \mathbf{a}_i^\top \lambda^*$. Then $\mathbf{a}_i^\top \mu = (1+\kappa)\alpha_i$. Combining (EC.10), (EC.11), and the regular-only baseline identity, it remains to prove the following fixed-product inequality for each $i \in [N]$:

$$\begin{aligned}(1+\kappa)\sum\nolimits_{k\in[q_i]} v_{i,k} + (2+\kappa)\sum\nolimits_{k\in[q_i]} \left(s - v_{i,k} - \alpha_i\right)^+ \\ \geq q_i\left(s + w - (1+\kappa)\alpha_i\right)^+.\end{aligned} \tag{EC.12}$$

Fix any product $i \in [N]$. For notational simplicity, we omit subscript i on v, q, and α. The second sum on the left-hand side of (EC.12) is nonnegative, so it suffices to prove

$$(1+\kappa)\sum\nolimits_{k\in[q]} \left[v_k + (s - v_k - \alpha)^+\right] \geq q\left(s + w - (1+\kappa)\alpha\right)^+. \tag{EC.13}$$

If $s + w - (1+\kappa)\alpha \leq 0$, the right-hand side of (EC.13) is zero. By the product-level decomposition, $\sum_{k\in[q]} v_k$ is the expected regular-only profit of this product, which is nonnegative under Assumption 1. The remaining term inside the brackets on the left-hand side is a sum of positive parts and is nonnegative. Thus, the left-hand side is nonnegative as well. Otherwise, the right-hand side equals $q\left(s + w - (1+\kappa)\alpha\right)$, and for every $k \in [q]$,

$$v_k + \left(s - v_k - \alpha\right)^+ \geq s - \alpha, \tag{EC.14}$$

since $(x)^+ \geq x$. Using (EC.14) and $\kappa s = w$, we obtain

$$\begin{aligned}
&(1+\kappa)\sum\nolimits_{k\in[q]}\left[v_k+(s-v_k-\alpha)^+\right]-q\big(s+w-(1+\kappa)\alpha\big)\\
&\quad\geq q\left[(1+\kappa)(s-\alpha)-\big(s+w-(1+\kappa)\alpha\big)\right]\\
&\quad= q(\kappa s-w)=0.
\end{aligned}$$

This proves (EC.13), completing the proof. □

EC.6. Proofs for Section 4

EC.6.1. Proof of Lemma 4

Proof. First record the rounded total selected by the construction. The total-sum guarantee of dependent rounding gives

$$\sum\nolimits_{i\in[N]}\hat{f}_i^{(t)}\in\left\{\left\lfloor\sum\nolimits_{i\in[N]}f_i^{(t)}\right\rfloor,\left\lceil\sum\nolimits_{i\in[N]}f_i^{(t)}\right\rceil\right\}.$$

Thus

$$z_{\cdot,\cdot}^{(t)}=\sum\nolimits_{i\in[N]}\lfloor u_{i,\cdot}^{(t)}\rfloor+\sum\nolimits_{i\in[N]}\hat{f}_i^{(t)}\in\left\{\left\lfloor\sum\nolimits_{i\in[N]}u_{i,\cdot}^{(t)}\right\rfloor,\left\lceil\sum\nolimits_{i\in[N]}u_{i,\cdot}^{(t)}\right\rceil\right\}.$$

By feasibility and the remaining-horizon capacity constraint,

$$\sum\nolimits_{i\in[N]}u_{i,\cdot}^{(t)}=\sum\nolimits_{(i,k)\in\mathcal{U}^{(t)}}u_{i,k}^{(t)}\leq(T-t+1)Y.$$

Since $(T-t+1)Y$ is an integer, $z_{\cdot,\cdot}^{(t)}\leq(T-t+1)Y$.

We next verify that the balanced load construction is well defined and respects capacity. Let $H_\tau \triangleq T-\tau+1$ and $M_\tau \triangleq m_\tau^{(t)}$. We claim by induction that $0\leq M_\tau\leq H_\tau Y$. This holds at $\tau=t$ by the bound above. If $M_\tau=0$, then $z_{\cdot,\tau}^{(t)}=0$. Otherwise,

$$1\leq z_{\cdot,\tau}^{(t)}=\left\lceil\frac{M_\tau}{H_\tau}\right\rceil\leq\min\{M_\tau,Y\}.$$

Hence $M_{\tau+1}=M_\tau-z_{\cdot,\tau}^{(t)}\geq 0$. Moreover,

$$M_{\tau+1}\leq M_\tau-\frac{M_\tau}{H_\tau}\leq(H_\tau-1)Y=H_{\tau+1}Y.$$

This proves the induction. At the final period $H_T=1$, the formula gives $z_{\cdot,T}^{(t)}=M_T$, so $m_{T+1}^{(t)}=0$. Therefore each period load is at most Y, and the loads contain exactly $z_{\cdot,\cdot}^{(t)}$ slots.

Because the total number of slots equals $\sum_i z_{i,\cdot}^{(t)}$, the slots can be filled with product labels having the prescribed product counts. This defines nonnegative integer counts $z_{i,\tau}^{(t)}$ with $\sum_{\tau=t}^{T}z_{i,\tau}^{(t)}=z_{i,\cdot}^{(t)}$ for every product i.

We now verify the four properties in Definition 3.

For (a), note that the construction assigns exactly $z_{\cdot,\tau}^{(t)}$ product labels to period τ, with $z_{\cdot,\tau}^{(t)}\leq Y$. Hence

$$\sum\nolimits_{i\in[N]}z_{i,\tau}^{(t)}\leq Y,\qquad\tau=t,\ldots,T.$$

For (b), the construction sets $z_{i,\cdot}^{(t)} = \lfloor u_{i,\cdot}^{(t)} \rfloor + \hat{f}_i^{(t)}$. Since $u_{i,\cdot}^{(t)} \le \tilde{q}_i^{(t-1)}$ and $\tilde{q}_i^{(t-1)}$ is integer, we have $z_{i,\cdot}^{(t)} \le \tilde{q}_i^{(t-1)}$.

For (c), fix a product i and write $u_{i,\cdot}^{(t)} = \lfloor u_{i,\cdot}^{(t)} \rfloor + f_i^{(t)}$ with $f_i^{(t)} \in [0,1)$. Since $u^{(t)}$ is in suffix form, the last $\lfloor u_{i,\cdot}^{(t)} \rfloor$ residual units of product i have reservation probability one, the preceding boundary unit has reservation probability $f_i^{(t)}$ if $f_i^{(t)} > 0$, and all earlier residual units have reservation probability zero. By dependent rounding,

$$\mathbb{E}\left[\hat{f}_i^{(t)} \,\middle|\, S^{(t-1)}\right] = f_i^{(t)}, \qquad i \in [N].$$

Therefore selecting the last $z_{i,\cdot}^{(t)} = \lfloor u_{i,\cdot}^{(t)} \rfloor + \hat{f}_i^{(t)}$ residual units gives

$$\mathbb{E}\left[I_{i,k}^{(t)} \,\middle|\, S^{(t-1)}\right] = u_{i,k}^{(t)}, \qquad (i,k) \in \mathcal{U}^{(t)}.$$

Finally, for (d), fix any cutoff period s. The number of product-i units scheduled in periods $s, \dots, T$ is $\sum_{\tau=s}^{T} z_{i,\tau}^{(t)}$. By the canonical suffix rule, these are the last $\sum_{\tau=s}^{T} z_{i,\tau}^{(t)}$ residual units of product i, and therefore form a suffix of $[\tilde{q}_i^{(t-1)}]$. □

EC.6.2. Proof of Lemma 5

Proof. By (d), the unexecuted tail consists, for each product i, of the last $\sum_{\tau=t+1}^{T} z_{i,\tau}^{(t)}$ residual units. By construction, these units are withheld from the regular channel in period t. Since the canonical suffix rule assigns earlier periods earlier units within the selected suffix, the units protected for current-period bag offers are disjoint from the tail, so the tail units remain physically available in the next residual state. The remaining-period bag capacities hold by (a):

$$\sum\nolimits_{i \in [N]} z_{i,\tau}^{(t)} \le Y, \qquad \tau = t+1, \dots, T.$$

More explicitly, period-t regular sales are at most $\tilde{q}_i^{(t-1)} - z_{i,\cdot}^{(t)}$, and hence

$$\tilde{q}_i^{(t)} \ge z_{i,\cdot}^{(t)} - z_{i,t}^{(t)} = \sum_{\tau=t+1}^{T} z_{i,\tau}^{(t)}.$$

By the canonical suffix rule, after the period-t protected units are removed, regular sales occur only outside the tail, and surviving units are re-indexed, the tail-scheduled units form a suffix for each product. Thus the tail induces an integral residual plan.

It remains to check the residual soft rows. The canonical suffix rule assigns the current-period counts and the unexecuted tail counts to disjoint portions of the same selected suffix; their union is exactly the set of residual units with $I_{i,k}^{(t)} = 1$. Therefore, for each row ℓ,

$$\sigma_{\ell,t}^{(t)}(Z^{(t)}) + \sigma_{\ell,>t}^{(t)}(Z^{(t)}) = \sum\nolimits_{(i,k) \in \mathcal{U}^{(t)}} a_{i,\ell} I_{i,k}^{(t)}.$$

Taking conditional expectations and using (c),

$$\mathbb{E}\left[\sigma_{\ell,t}^{(t)}(Z^{(t)})+\sigma_{\ell,>t}^{(t)}(Z^{(t)})\,\middle|\,S^{(t-1)}\right]=\sum\nolimits_{(i,k)\in\mathcal{U}^{(t)}} a_{i,\ell}u_{i,k}^{*(t)}.$$

Since $u^{*(t)}$ is feasible for the residual LP, the right hand side is at most $\tilde{b}_\ell^{(t-1)}$, and hence $\mathsf{Slack}_\ell^{(t)}\geq 0$. Therefore,

$$\sigma_{\ell,>t}^{(t)}(Z^{(t)})\leq\sigma_{\ell,>t}^{(t)}(Z^{(t)})+\mathsf{Slack}_\ell^{(t)}=\tilde{b}_\ell^{(t)}.$$

Thus the unexecuted tail is feasible for the next residual LP. □

EC.6.3. Proof of Proposition 1

Proof. We first show that the policy is well defined. The initial residual LP is feasible because $b_\ell\geq 0$ and the zero plan is feasible. Suppose the residual LP is feasible at the beginning of period $t<T$. By Lemma 5, after period t is executed, the unexecuted tail of the admissible rounded schedule induces a feasible integral solution to the next residual LP. Hence the next residual LP is feasible. By induction, the residual LP is feasible whenever the policy solves it.

We next prove the hard constraints. For the regular-display constraints, the policy displays

$$x_{i,t}=\min\left\{X_i,\tilde{q}_i^{(t-1)}-z_{i,\cdot}^{(t)}\right\}.$$

By (b), $0\leq z_{i,\cdot}^{(t)}\leq\tilde{q}_i^{(t-1)}$. Hence we have $x_{i,t}\leq\tilde{q}_i^{(t-1)}\leq q_{i,t-1}$, so the inventory constraint holds pathwise. The same definition gives $x_{i,t}\leq X_i$, so the regular-sale cap constraint holds pathwise.

For the per-period bag-cap constraint, (a) gives

$$y_t=\sum\nolimits_{i\in[N]} z_{i,t}^{(t)}\leq Y,$$

so bag-display cap constraint holds pathwise.

It remains to verify the two fulfillment constraints. At the end of the horizon, the policy sets

$$z_i=\sum\nolimits_{t\in[T]} z_{i,t}^{(t)},\qquad i\in[N].$$

Together with $y_t=\sum_{i\in[N]} z_{i,t}^{(t)}$, this gives

$$\sum\nolimits_{i\in[N]} z_i=\sum\nolimits_{t\in[T]} y_t,$$

so the policy allocates exactly one unit to each accepted bag. Moreover, the $z_{i,t}^{(t)}$ units protected for period-t bag offers are removed from the residual inventory and are never pooled, displayed, or sold later. Hence they remain available for final fulfillment, so

$$z_i\leq q_{i,T},\qquad i\in[N],$$

which establishes productwise inventory availability for final fulfillment.

Lastly, we prove the soft constraints. Fix a row ℓ. By (4),

$$\mathbb{E}\left[\tilde{b}_\ell^{(t)} \,\middle|\, S^{(t-1)}\right] = \tilde{b}_\ell^{(t-1)} - \mathbb{E}\left[\sigma_{\ell,t}^{(t)}(Z^{(t)}) \,\middle|\, S^{(t-1)}\right].$$

Taking expectations and summing over $t = 1, \ldots, T$, with $\tilde{b}_\ell^{(0)} = b_\ell$, gives

$$\sum\nolimits_{t=1}^{T} \mathbb{E}\left[\sigma_{\ell,t}^{(t)}(Z^{(t)})\right] = b_\ell - \mathbb{E}\left[\tilde{b}_\ell^{(T)}\right].$$

After the last update, the scheduled tail is empty, so $\tilde{b}_\ell^{(T)} = \mathsf{Slack}_\ell^{(T)} \geq 0$. Hence

$$\sum\nolimits_{t=1}^{T} \mathbb{E}\left[\sigma_{\ell,t}^{(t)}(Z^{(t)})\right] \leq b_\ell.$$

Since

$$\sigma_{\ell,t}^{(t)}(Z^{(t)}) = \sum\nolimits_{i\in[N]} a_{i,\ell} z_{i,t}^{(t)},$$

we obtain

$$\sum\nolimits_{i\in[N]} a_{i,\ell} \mathbb{E}\left[\sum\nolimits_{t=1}^{T} z_{i,t}^{(t)}\right] \leq b_\ell.$$

Thus (1) holds for row ℓ, and hence for all $\ell \in [L]$. □

EC.6.4. Proof of Lemma 6

Proof. Fix $S^{(t-1)}$, and let $u^{*(t)}$ be a suffix-form optimal solution of the residual LP. Let $Z^{(t)}$ be the admissible rounded schedule drawn from $u^{*(t)}$.

After $Z^{(t)}$ is drawn and period t is executed, let $C_{t+1}\big(S^{(t)}; Z^{(t)}\big)$ denote the expected value, conditional on the realized next residual state $S^{(t)}$ and on the unexecuted tail of $Z^{(t)}$, of committing the tail units to future bags and running the remaining residual units greedily through the regular channel from period $t+1$ onward. This continuation value is a random variable because it depends on the realized next state. When $t = T$, the tail is empty and this continuation value is zero.

First consider the fixed continuation that executes the current-period part of $Z^{(t)}$ and then follows the unexecuted tail of the same schedule. Conditional on $S^{(t-1)}$ and $Z^{(t)}$, this continuation assigns residual unit (i,k) to the surprise-bag channel if $I_{i,k}^{(t)} = 1$: units in the current component are protected for period-t bag offers, and units in the tail are used in the fixed continuation. All other residual units are left for the regular channel. The period-t profit P_t contains the regular sales and bag rewards realized in period t. The tail-continuation value contains only the expected future contribution from periods $t+1, \ldots, T$, including future bag rewards and the terminal waste cost on residual units not assigned to bags. Thus their sum is the value of this fixed suffix reservation in the residual instance. The canonical suffix rule, together with (a) and (b), provides the required suffix form and schedule feasibility. The unselected prefix is run greedily through the regular channel, so each unselected unit (i,k) contributes its residual regular value $v_{i,k}^{(t)}$ in expectation,

while each selected unit contributes s. Because the residual state carries the conditional distribution $\mathcal{D}^{(t)}(\cdot \mid \boldsymbol{d}_{1:t-1})$, these residual values are exactly the conditional expected contributions given the current state and schedule. The random seed used in period t to draw $Z^{(t)}$ is independent of the demand process and the policy's preceding random seeds, and hence of remaining demand conditional on $S^{(t-1)}$. Hence

$$\mathbb{E}\left[P_t + C_{t+1}\left(S^{(t)}; Z^{(t)}\right) \mid S^{(t-1)}, Z^{(t)}\right] = \sum\nolimits_{(i,k)\in\mathcal{U}^{(t)}} v_{i,k}^{(t)} + \sum\nolimits_{(i,k)\in\mathcal{U}^{(t)}} \left(s - v_{i,k}^{(t)}\right) I_{i,k}^{(t)}.$$

Taking conditional expectation over the randomization used in period t to draw $Z^{(t)}$ and using (c),

$$\mathbb{E}\left[I_{i,k}^{(t)} \,\middle|\, S^{(t-1)}\right] = u_{i,k}^{*(t)}, \qquad (i,k)\in\mathcal{U}^{(t)},$$

we obtain

$$\mathbb{E}\left[P_t + C_{t+1}\left(S^{(t)}; Z^{(t)}\right) \mid S^{(t-1)}\right] = \sum\nolimits_{(i,k)\in\mathcal{U}^{(t)}} v_{i,k}^{(t)} + \sum\nolimits_{(i,k)\in\mathcal{U}^{(t)}} \left(s - v_{i,k}^{(t)}\right) u_{i,k}^{*(t)}.$$

Since $u^{*(t)}$ is optimal for the residual LP, the right-hand side is exactly $\mathrm{LP}_t(S^{(t-1)})$.

If $t = T$, the tail-continuation value is zero, so the terminal convention immediately gives the claimed inequality. Suppose henceforth that $t < T$.

On every realized branch, Lemma 5 shows that the unexecuted tail of $Z^{(t)}$ induces a feasible integral solution to the residual LP at $S^{(t)}$. In particular, the induced tail solution has suffix form and can be offered and fulfilled over periods $t+1,\dots,T$ within the remaining bag-offer cap. Applying the same unit-value accounting to this next residual instance, the objective value of the induced tail solution is $C_{t+1}\left(S^{(t)}; Z^{(t)}\right)$. Hence, by optimality of the residual LP at the next residual state,

$$\mathrm{LP}_{t+1}(S^{(t)}) \geq C_{t+1}\left(S^{(t)}; Z^{(t)}\right) \qquad \text{almost surely.}$$

Combining this inequality with the previous display proves the claim. □

EC.7. Proof of Theorem 3

Throughout this section, we work in the symmetric uniform-demand setting of Section 5.1: $N, T \geq 2$, $X_i = X \geq 1$, $q_i = TX$, $r_i = r$ for all $i \in [N]$, $Y \geq 1$, and there are no soft bag-composition constraints. We assume $r > s > 0$ and $s \geq w \geq 0$, with $d_{i,t}$ i.i.d. $\mathrm{Uniform}\{0,1,\dots,D\}$. We call periods $1,\dots,T-1$ prefinal.

Proof strategy. Let DA^* be a DA-optimal policy. We will select an IA-optimal policy IA^* and construct its DA shadow policy. The shadow is DA-feasible and has the same profit as IA^* on every demand path. To prove the theorem, it is therefore enough to construct a DA-feasible policy DA' satisfying

$$\mathbb{E}\left[\textsc{Profit}\left[\mathrm{DA}'; I\right]\right] > \mathbb{E}\left[\textsc{Profit}[\mathrm{IA}^*; I]\right].$$

The remainder of the proof constructs two complementary DA policy witnesses: one reallocates an existing bag commitment inherited from IA, and the other creates an additional period-1 bag commitment. The optimality of DA^* then proves the theorem.

Before constructing the DA shadow policy, we select an IA-optimal representative that displays every available unit not assigned to a surprise bag, up to the display cap. The next lemma shows that this full-display property is without loss of optimality. We now introduce the IA notation needed to state it precisely.

Under IA, a product unit is immediately assigned to each accepted surprise bag. For each product i and period t, let $\tilde{q}_i^{(t-1)}$ denote the units available at the period-t decision, after regular sales and bag assignments in periods $1,\dots,t-1$. Let $z_{i,t}$ be the units of product i assigned to bags accepted in period t, and let $x_{i,t}$ be the units displayed through the regular channel. IA feasibility requires

$$z_{i,t} + x_{i,t} \le \tilde{q}_i^{(t-1)}, \qquad i \in [N], \quad t \in [T].$$

LEMMA EC.3. *There exists a deterministic Markov IA-optimal policy* IA^*. *Its period-*t *decision depends only on the IA state* $(t, (\tilde{q}_i^{(t-1)})_{i\in[N]})$. *Moreover, on every realized demand path,*

$$x_{i,t} = \min\left\{X,\, \tilde{q}_i^{(t-1)} - z_{i,t}\right\}, \qquad i \in [N], \quad t \in [T]. \tag{EC.15}$$

The proof is deferred to Section EC.7.1.

DEFINITION EC.1 (DA-SHADOW OPERATOR). Define the *DA-shadow operator* $\mathcal{DA}(\cdot)$ on the class of nonanticipative IA-feasible policies as follows. For any such policy π, couple π and $\mathcal{DA}(\pi)$ through the same realization of π's internal randomness. In each period, $\mathcal{DA}(\pi)$ reproduces the bag-acceptance and regular-display decisions that π would make after the same preceding demand history and corresponding preceding IA decisions, but postpones each product-unit assignment until the end of the horizon. At the horizon, it fulfills each bag with the product unit that π assigned to that bag.

LEMMA EC.4. *For every nonanticipative IA-feasible policy* π, $\mathcal{DA}(\pi)$ *is a nonanticipative DA-feasible policy and has the same profit as* π *on every joint realization of demands and* π*'s internal randomness.*

The proof is deferred to Section EC.7.1.

Fix the deterministic Markov full-display IA-optimal representative IA^* supplied by Lemma EC.3. For $t \in [T]$, write its uncommitted residual-inventory vector as

$$\tilde{q}^{(t-1)} \triangleq (\tilde{q}_i^{(t-1)})_{i\in[N]}.$$

In each period, $\mathcal{DA}(\mathrm{IA}^*)$ accepts the same number, $\sum_{i\in[N]} z_{i,t}$, of bags and displays $x_{i,t}$ units of each product i; it postpones the corresponding product-unit assignments until the end of the horizon. For each product i and $t \in \{0,\dots,T\}$, the shadow records the cumulative IA bag assignments through period t as

$$Z_{i,t} := \sum_{u=1}^{t} z_{i,u}, \qquad Z_{i,0} := 0, \qquad i \in [N].$$

Let $d_{i,t}$ denote regular demand for product i in period t. To reproduce the IA decision sequence, the shadow tracks the remaining inventory after IA bag assignments and regular sales:

$$\tilde{q}_i^{(t)} := \tilde{q}_i^{(t-1)} - z_{i,t} - \min\{d_{i,t}, x_{i,t}\}, \qquad i \in [N], \quad t \in [T].$$

The next lemma formalizes this reassignment. If a product unit assigned by IA to a prefinal bag can instead be displayed in period T, while a unit of another product can fulfill that bag, it constructs a feasible DA policy and identifies when that policy strictly improves upon the DA shadow.

LEMMA EC.5. *Consider $\mathcal{DA}(\mathrm{IA}^*)$, the DA shadow defined in Definition EC.1 for the full-display IA-optimal policy IA^*. Suppose the DA shadow reaches a prefinal state with positive probability. At that state, let $(\tilde{q}^{(T-1)}, Z_{T-1})$ denote its uncommitted residual inventory and cumulative IA bag assignments. Let (z_T, x_T) denote the period-T shadow decision at that state. Suppose there exists a product j such that*

$$Z_{j,T-1} \geq 1, \qquad x_{j,T} < X. \tag{EC.16}$$

Thus, the shadow's recorded bag assignment includes a product-j unit inherited from a prefinal IA bag assignment, while product j is under-displayed in period T. Suppose that at least one of the following feasible and strictly profitable reassignments is available:

(a) ***Idle-unit reassignment.*** *There exists a product $k \neq j$ such that*

$$\tilde{q}_k^{(T-1)} - z_{k,T} - x_{k,T} \geq 1, \tag{EC.17}$$

and the resulting expected profit gain is strictly positive:

$$(r+w)\boldsymbol{Pr}[d_{j,T} \geq x_{j,T} + 1] > 0. \tag{EC.18}$$

(b) ***Bag-replacement reassignment.*** *There exists a product $k \neq j$ such that*

$$z_{k,T} \geq 1, \tag{EC.19}$$

and the resulting expected profit gain is strictly positive:

$$(r+w)\boldsymbol{Pr}[d_{j,T} \geq x_{j,T} + 1] - (s+w) > 0. \tag{EC.20}$$

Then there exists a nonanticipative DA-feasible policy that revises the product assignment of a previously accepted bag and uses the resulting flexibility to create an additional period-T display. Its expected profit is strictly larger than that of the shadow.

Proof. Both constructions use the same local reassignment. If the specified prefinal state is reached, the policy uses one product-j unit assigned by IA to a prefinal bag for an additional period-T display. At the horizon, the selected prefinal bag receives a product-k unit instead; every other recorded bag assignment follows the shadow. For either reassignment, choose a product k satisfying its donor condition. The reassignments differ only in the source of that replacement unit. In the idle-unit reassignment, the product-k unit is idle. In the bag-replacement reassignment, it would fulfill a bag accepted in period T, and the policy declines that bag at its period-T decision.

Feasibility. In both constructions, the modified policy is feasible. Because the shadow postpones the product assignment for the selected prefinal bag, its product-j unit can be reassigned. Moreover, because the shadow is derived from the full-display IA-optimal policy IA^* and $x_{j,T} < X$, the full-display property implies

$$x_{j,T} = \tilde{q}_j^{(T-1)} - z_{j,T},$$

Thus every product-j unit remaining after the period-T bag decision is displayed. The reassignment therefore creates exactly one additional display opportunity and one unit of bag-assignment shortfall. An idle product-k unit exists by (EC.17) and supplies the replacement in the idle-unit reassignment. In the bag-replacement reassignment, (EC.19) identifies a product-k unit assigned to a period-T bag. Declining that bag eliminates the need for this unit, which can therefore replace the reclaimed product-j unit. All other recorded bag assignments and displays remain unchanged.

Profit gain. Relative to the shadow, the modification reallocates one product-j unit from a prefinal bag to the period-T display and replaces it with a product-k unit. The additional display sells precisely when

$$d_{j,T} \geq x_{j,T} + 1.$$

Because period-T demand is independent of reaching that state, the conditional probability of this event equals its unconditional probability. Under idle-unit reassignment, the donor product-k unit would otherwise be wasted. If the additional display sells, the modification earns one regular-sale reward and avoids one unit of waste, for a gain of $r + w$. If the additional display does not sell, one unit is still wasted, now from product j rather than product k, so profit is unchanged. Therefore, the conditional expected gain equals

$$(r + w)\mathrm{Pr}[d_{j,T} \geq x_{j,T} + 1]\,.$$

Under bag-replacement reassignment, the period-T bag is declined. If the additional display sells, the modification gains one regular-sale reward and loses one bag reward, for a gain of $r - s$. If the additional display does not sell, the reclaimed product-j unit is wasted and the bag reward is lost, for a gain of $-s - w$. The conditional expected gain equals

$$\mathrm{Pr}[d_{j,T} \geq x_{j,T} + 1]\,(r - s) + \Big(1 - \mathrm{Pr}[d_{j,T} \geq x_{j,T} + 1]\Big)(-s - w),$$

which simplifies to

$$(r+w)\mathrm{Pr}[d_{j,T} \geq x_{j,T}+1] - (s+w).$$

Using a fixed tie-breaking rule, the policy determines the prefinal state, the product k, and the modified period-T display from demands observed through period $T-1$ and the simulated deterministic Markov shadow decision. It makes the modified bag assignment only after observing period-T demand. Thus the policy is nonanticipative. In each reassignment, the policy agrees with the shadow unless it reaches the specified state, so its profit difference from the shadow is zero otherwise. Because that state has positive probability, a strictly positive conditional expected gain given that state yields a strictly positive unconditional expected gain. The respective positivity conditions (EC.18) and (EC.20) therefore imply that the constructed policy has strictly larger expected profit than the shadow. □

We call the prefinal demand history in which every product's demand equals X in every prefinal period the full-sales reference history.

Lemma EC.5 exploits flexibility in an existing bag commitment. We now consider the complementary case in which the IA-optimal policy accepts no bag before the last period on the full-sales reference history. The next lemma gives a DA policy witness that exploits this unused early bag capacity. It accepts one additional period-1 bag and postpones its product assignment. While this additional commitment remains pending, the policy otherwise follows the DA shadow. The commitment is eventually resolved either by replacing an omitted later shadow bag, by inventory surplus that guarantees fulfillment, or, if neither occurs, through a final-period fallback.

Under the demand condition $D \geq X$ used in the next lemma, set

$$\zeta := \frac{X}{D-X+1}, \qquad p_0 := (1+\zeta)^{-N}, \qquad v_T^{\mathrm{marg}} := \frac{r-w\zeta}{1+\zeta}. \tag{EC.21}$$

Here ζ is the one-product odds of a positive full-display leftover relative to a full sellout, p_0 is the corresponding probability that all products sell their full displays in one period, and v_T^{marg} is the expected marginal profit contribution of the last unit of a full final-period display.

LEMMA EC.6. *Consider the DA shadow $\mathcal{DA}(\mathrm{IA}^*)$ of the deterministic Markov full-display IA-optimal policy IA^* fixed above. Suppose that IA^* accepts no bag before the last period on the full-sales reference history. Assume $Y \leq (N-1)X$ and $D \geq X$. Then there exists a nonanticipative DA-feasible policy π^{add} that accepts one additional period-1 bag and postpones its product assignment. With ζ, p_0, and v_T^{marg} defined in (EC.21), and $S(\cdot)$ defined in (5), its expected profit gain relative to the shadow satisfies*

$$\mathbb{E}\left[\textsc{Profit}\left[\pi^{\mathrm{add}}; I\right] - \textsc{Profit}[\mathcal{DA}(\mathrm{IA}^*); I]\right] \geq p_0^{T-1}\left[(s+w)S(\zeta) - \left(v_T^{\mathrm{marg}} - s\right)^+\right]. \tag{EC.22}$$

Moreover, if $D < D_{\mathrm{U}}$, with D_{U} defined in (8), then this expected gain is strictly positive.

Proof. We first construct the policy and verify its feasibility. We then identify the inventory-surplus events that secure the additional commitment and bound its expected gain.

Policy construction and feasibility. Because period 1 occurs before any demand is observed and the initial state is deterministic, IA^* accepts no bag in period 1 on any demand path. Since $Y \geq 1$, π^{add} can accept one additional period-1 bag. This additional commitment is never cancelled. Until it is replaced or secured, call the additional commitment *pending*. The pending commitment is resolved only in one of three ways: replacement by a later shadow bag, securing through inventory surplus, or the final-period fallback. While it is pending, π^{add} copies every shadow display, except for the one-unit reduction in the final-period fallback. If the shadow accepts at least one bag in a later period, π^{add} accepts one fewer bag in that period. The period-1 bag then replaces the omitted shadow bag: at the horizon, it receives the product unit that the shadow would have assigned to the omitted bag, while all other shadow assignments are unchanged. The policy thereafter copies every shadow decision.

Before such a replacement, no shadow bag has been accepted in an earlier period. At the beginning of every pending period, the physical inventory of π^{add} therefore equals the uncommitted residual inventory $\tilde{q}^{(t-1)}$, so every copied shadow display is feasible. Indeed, any shadow bag accepted while the commitment is pending would replace it immediately. Therefore, if the commitment remains pending after the period-t bag decision for $t < T$, then $z_{i,t} = 0$ for every product i, and

$$\tilde{q}_i^{(t-1)} \geq (T - t + 1)X, \qquad i \in [N].$$

The full-display property then gives $x_{i,t} = X$. Even if every such display sells, at least $(T-t)X$ units of every product remain afterward. Hence no prefinal display reduction is required while the commitment is pending.

If the commitment reaches period T still pending, a period-T shadow bag again replaces it if one is accepted. Otherwise, π^{add} copies the shadow display whenever

$$\sum\nolimits_i \left(\tilde{q}_i^{(T-1)} - x_{i,T}\right) \geq 1,$$

using an undisplayed unit for the additional bag. In the remaining case, no period-T shadow bag is accepted, so $z_{i,T} = 0$ for every product i. Because the commitment is pending, no shadow bag was accepted earlier either, and the same no-prior-bag argument gives $\tilde{q}_i^{(T-1)} \geq X$. The full-display property therefore gives $x_{i,T} = X$. The failed residual-capacity condition makes this residual-capacity sum equal to zero; since each summand is nonnegative,

$$\tilde{q}_i^{(T-1)} = x_{i,T} = X, \qquad i \in [N].$$

The policy withholds one unit from the smallest-index shadow display and reserves it for the additional bag. Thus the construction is DA-feasible on every demand path.

Pathwise profit comparison. The policy construction above has three economically distinct outcomes. If a later shadow bag replaces the additional commitment, the modification leaves profit unchanged. If the additional commitment is fulfilled in addition to all shadow decisions, the policy earns one additional bag reward and avoids one unit of waste, for a gain of $s+w$. The remaining possibility is that the commitment must ultimately be fulfilled through the period-T fallback, which may sacrifice one regular-sale opportunity. We will show below that a negative gain can arise only on the all-zero-leftover prefinal history.

To lower-bound the expected gain, it therefore suffices to identify a tractable collection of paths on which fulfillment in addition to all shadow decisions is guaranteed.

A tractable family of favorable surplus paths. We formalize the inventory-surplus resolution through the following certification condition. After period-u demand while the commitment remains pending, declare it *secured* if

$$\sum\nolimits_i \tilde{q}_i^{(u)} > N(T-u)X + (T-u)Y. \tag{EC.23}$$

Once secured, π^{add} copies every subsequent shadow decision. Over the remaining horizon, regular sales can consume at most $N(T-u)X$ units, while future shadow bags can require at most $(T-u)Y$ units for fulfillment at the horizon. The strict inequality therefore leaves at least one inventory unit beyond all possible future shadow uses. That unit fulfills the additional bag, yielding profit gain $s+w$ relative to the shadow.

We next identify disjoint surplus events satisfying this condition.

If all X units of every product are displayed in period t, let

$$K_t := \sum\nolimits_{i=1}^{N} \left(X - \min\{d_{i,t}, X\}\right), \qquad \Pr[K_t = 0] = p_0.$$

Conditioning on the number and leftover amounts of products with positive leftover yields

$$\frac{\Pr[K_t > b]}{p_0} = \sum\nolimits_{j=1}^{N} \binom{N}{j} \sum_{\substack{n_1,\dots,n_X \geq 0 \\ n_1+\cdots+n_X=j \\ \sum_{r=1}^{X} r n_r > b}} \binom{j}{n_1,\dots,n_X} \left(\frac{\zeta}{X}\right)^j. \tag{EC.24}$$

For each m, (EC.24) identifies the odds-normalized probability of more than mY units of full-display leftover. Substituting these probabilities into the corresponding terms of (5) gives

$$S(\zeta) = \sum\nolimits_{m=1}^{T-1} p_0^{-m} \Pr[K_t > mY]. \tag{EC.25}$$

For $m = 1, \dots, T-1$, define

$$E_m := \{K_1 = \cdots = K_{T-m-1} = 0,\ K_{T-m} > mY\}.$$

For $m = T-1$, the zero-leftover prefix is empty. These events are disjoint. On E_m, its zero-leftover prefix reproduces the reference predecision states, so the Markov policy accepts no shadow bag through period

$T-m$. The additional commitment is therefore pending when period-$T-m$ demand is observed. After that period,

$$\sum_i \tilde{q}_i^{(T-m)} = NmX + K_{T-m} > NmX + mY,$$

so (EC.23) secures the commitment and yields gain $s+w$. By independence across periods,

$$\sum_{m=1}^{T-1} \Pr[E_m] = p_0^{T-1} S(\zeta)\,. \tag{EC.26}$$

Expected-gain bound. Let Δ^{add} be the realized profit difference from the shadow. The certification events contribute $s+w$ each. If a later shadow bag is accepted while the commitment is pending, the additional commitment replaces that bag, and the gain is zero. Consider every remaining path outside the all-zero-leftover prefinal history. Let $\sigma < T$ be the first period with $K_\sigma > 0$. The preceding zero-leftover periods reproduce the reference predecision states, so the shadow accepts no bag through period σ. Because $K_\sigma > 0$, some product i has more than $(T-\sigma)X$ units after period σ. If no later shadow bag replaces the commitment, the remaining regular displays can sell at most $(T-\sigma)X$ units of that product. Hence at least one such unit survives for the additional bag, yielding gain $s+w$.

It remains to consider the all-zero-leftover prefinal history, whose probability is p_0^{T-1}. If the shadow accepts a period-T bag, the additional commitment replaces it and the gain is zero. Otherwise, the final-period fallback withholds one displayed unit. Its expected relative gain is $s - v_T^{\mathrm{marg}}$, so the conditional expected loss is at most

$$\left(v_T^{\mathrm{marg}} - s\right)^+.$$

Combining this possible loss with the favorable certification events gives

$$\begin{aligned}\mathbb{E}\left[\Delta^{\mathrm{add}}\right] &\geq (s+w)p_0^{T-1}S(\zeta) - p_0^{T-1}\left(v_T^{\mathrm{marg}} - s\right)^+ \\ &= p_0^{T-1}\left[(s+w)S(\zeta) - \left(v_T^{\mathrm{marg}} - s\right)^+\right],\end{aligned}$$

which is (EC.22).

The capacity condition gives $Y < NX$, so $S(\zeta) > 0$. If $v_T^{\mathrm{marg}} \leq s$, the lower bound is strictly positive. Otherwise,

$$v_T^{\mathrm{marg}} - s = \frac{(r-s) - (s+w)\zeta}{1+\zeta},$$

and positivity is equivalent to

$$\zeta + (1+\zeta)S(\zeta) > \frac{r-s}{s+w}.$$

The left-hand side is strictly increasing as a function of $\zeta > 0$. Thus $D < D_{\mathrm{U}}$ implies the required inequality by (6). The policy uses only observed demands and the simulated Markov shadow state, so it is nonanticipative. □

Completion of the proof of Theorem 3. We now combine Lemmas EC.5 and EC.6 to prove the theorem. We first verify that the two demand thresholds are ordered, and then show that one of the two policy witnesses applies according to the IA-optimal policy's behavior on the full-sales reference history.

We first verify that the two demand thresholds are ordered. The capacity bound $Y \le (N-1)X$ implies $Y < NX$. Hence $S(y) > 0$ for every $y > 0$. Moreover, $S(y)$ is a polynomial with nonnegative coefficients and no constant term. Therefore

$$f(y) := y + (1+y)S(y)$$

is continuous and strictly increasing from zero to infinity, so $\phi(J)$ exists and is unique.

Evaluating f at $(r-s)/(s+w)$ gives a value strictly larger than the right-hand side of (6). Thus

$$\phi(J) < \frac{r-s}{s+w},$$

and consequently

$$D_{\mathrm{U}} > \frac{(s+w)X}{r-s} + X - 1 = \frac{(r+w)X}{r-s} - 1 \ge D_{\mathrm{L}}.$$

This proves $D_{\mathrm{L}} < D_{\mathrm{U}}$.

Now fix an integer $D \in (D_{\mathrm{L}}, D_{\mathrm{U}})$. The lower bound implies $D \ge X$.

Now consider the full-sales reference history defined above. It has positive probability. Along this history, the selected IA-optimal policy either accepts a bag before the final period, creating an early commitment that DA may reallocate, or it does not, leaving early bag capacity unused. The early-acceptance alternative yields a reassignment policy witness (Lemma EC.5), whereas the no-prefinal-bag alternative yields an additional-bag policy witness (Lemma EC.6).

Early bag acceptance on the reference history.

Let $\tau \le T-1$ be the first such period, and choose a product i from which IA* assigns at least one unit to a bag in that period. Tag one such bag. Keep the reference history through period $\tau - 1$, and from periods τ through $T-1$ take demand X for product i and zero for every product $k \ne i$. This prefinal history has positive probability.

Because no bag is accepted before τ, each product enters that period with $(T-\tau+1)X$ units. The donor products $k \ne i$ make no subsequent regular sales before the last period. At most $(T-\tau)Y - 1$ bags are assigned using donor-product units, because a product-i unit is assigned to the tagged bag. If $\tilde{q}^*$ denotes the IA inventory entering the last period, then

$$\sum\nolimits_{k \ne i} (\tilde{q}^*_k - X) \ge (T-\tau)\big((N-1)X - Y\big) + 1. \tag{EC.27}$$

Product i, meanwhile, sells every displayed unit and loses the tagged unit to a bag. Write $L := T - \tau$. It enters period τ with $(L+1)X$ units. In each of the L remaining prefinal periods, full display and demand X give the recursion

$$\tilde{q}_i^{(u)} = \max\left\{\tilde{q}_i^{(u-1)} - b_u - X, 0\right\}, \qquad u = \tau, \dots, T-1,$$

where $b_u := z_{i,u}$ is the period-u product-i bag assignment. Iterating this recursion gives

$$\tilde{q}_i^* = \max\left\{X - \sum\nolimits_{u=\tau}^{T-1} b_u, 0\right\} \le X - 1,$$

because $b_\tau \ge 1$. Let Z_i^* be the cumulative IA assignment of product i before the last period. The resulting prefinal shadow state, with uncommitted residual inventory $\tilde{q}^*$ and cumulative IA assignments Z^*, has positive probability. If $x_{i,T}^*$ is its last-period IA display, the tagged prefinal bag gives $Z_i^* \ge 1$, and the preceding inequality gives $x_{i,T}^* < X$. Thus DA can use the tagged product-i unit for an additional regular display, provided another product can cover the recorded bag assignment.

Suppose first that $2Y \le (N-1)X$. Then the right-hand side of (EC.27) exceeds Y. If no donor $k \ne i$ had an idle unit, then its full display would leave no unit outside its last-period bags and display. Thus

$$\tilde{q}_k^* - X \le z_{k,T}^*, \qquad k \ne i.$$

Consequently,

$$\sum\nolimits_{k \ne i} (\tilde{q}_k^* - X) \le \sum\nolimits_{k \ne i} z_{k,T}^* \le Y,$$

contradicting (EC.27). Hence some donor $k \ne i$ satisfies

$$\tilde{q}_k^* - z_{k,T}^* - x_{k,T}^* \ge 1.$$

Thus the donor condition in Reassignment (a) holds at this state with $j = i$.

Now suppose $2Y > (N-1)X$. The capacity bound $Y \le (N-1)X$ and (EC.27) imply that some donor $k \ne i$ has $\tilde{q}_k^* > X$. If this donor has a unit outside its last-period assignments to bags and display, then the donor condition in Reassignment (a) is satisfied at this state with $j = i$. Otherwise, full display and $\tilde{q}_k^* > X$ imply $z_{k,T}^* \ge 1$, so the donor condition in Reassignment (b) is satisfied at this state with $j = i$.

Thus at least one donor condition in Lemma EC.5 holds at this state. It remains to verify the corresponding gain criterion. For the gain comparison, let

$$\ell := x_{i,T}^* + 1 \le X.$$

This is the demand threshold for selling the additional display. In the idle-unit case, $\ell \le X \le D$ makes the expected-profit gain in Reassignment (a) strictly positive. In the period-T bag case, the expected-profit gain in Reassignment (b) is

$$(r+w)\Pr[d_{i,T} \ge \ell] - (s+w).$$

Because $\ell \le X$, this quantity is at least

$$\frac{r+w}{1+\zeta} - (s+w) = v_T^{\mathrm{marg}} - s > 0.$$

The final inequality follows from $D > D_{\mathrm{L}}$ in the present capacity regime. Thus Lemma EC.5 supplies a nonanticipative DA-feasible policy whose expected profit is strictly larger than that of the shadow. By Lemma EC.4 and DA optimality, this yields the desired strict inequality under early bag acceptance.

No early bag acceptance on the reference history.

The no-prefinal-bag alternative gives the corresponding condition of Lemma EC.6. The capacity bound gives $Y \le (N-1)X$, while $D > D_{\mathrm{L}} \ge X - 1$ and integrality imply $D \ge X$. Hence all hypotheses of Lemma EC.6 are satisfied. Because $D < D_{\mathrm{U}}$, that lemma supplies a nonanticipative DA-feasible policy with strictly larger expected profit than $\mathcal{DA}(\mathrm{IA}^*)$. By Lemma EC.4 and DA optimality, the desired strict inequality follows under no early bag acceptance.

Since these alternatives are exhaustive,

$$V^{\mathrm{DA}}(I) > V^{\mathrm{IA}}(I),$$

completing the proof. □

EC.7.1. Proofs of Supporting Lemmas

Proof of Lemma EC.3. The horizon, the feasible IA state and decision sets, and the demand support are finite. Because future demand is independent of observed history, backward induction yields a deterministic Markov IA-optimal policy. We show that this backward induction can select a full-display maximizing decision at every feasible period-t inventory vector.

For a period-t IA inventory vector $\mathbf{v}$, let $V_t(\mathbf{v})$ be the optimal expected continuation profit. We first show that, whenever $\mathbf{v}_i \ge 1$,

$$V_t(\mathbf{v}) - V_t(\mathbf{v} - e_i) \le r. \tag{EC.28}$$

Couple an optimal continuation from $\mathbf{v}$ with a continuation from $\mathbf{v} - e_i$, marking the extra product-i unit in the former system. The smaller system simulates the larger continuation until the marked unit is removed. Whenever a decision leaves a product-i unit unassigned and undisplayed, that same decision is feasible in the smaller system.

Otherwise, every product-i unit is assigned or displayed. If one is assigned to a bag, the smaller system accepts one fewer bag and removes the corresponding product-i assignment. It loses at most $s \le r$, after which the inventory states coincide. If none is assigned to a bag, every product-i unit is displayed. The smaller system displays one fewer. If that additional display sells, it loses r and the inventory states coincide; otherwise regular sales agree and the one-unit discrepancy persists. If the marked unit is never removed, it is

an additional unit discarded at the horizon and cannot increase profit. Thus the constructed nonanticipative continuation from $\mathbf{v} - e_i$ has expected profit at least $V_t(\mathbf{v}) - r$, proving (EC.28).

Fix a feasible period-t inventory vector and an optimal decision that violates (EC.15) for some product i. Then a product-i unit is unassigned and undisplayed. Keep the bag decision and all other displays unchanged, and add that unit to the display. If the added unit does not sell, the continuation state is unchanged. If it sells, it earns one additional regular-sale reward r and leaves one fewer product-i unit for the remaining periods, whose optimal continuation profit is lower by at most r by (EC.28). Thus the modification has expected profit no smaller than that of the original decision. Because the original decision is optimal, the two decisions must tie, so the modified decision is also optimal. Repeating this exchange until (EC.15) holds for every product produces a full-display optimal decision at the state. Applying this construction at every feasible state yields the claimed policy. □

Proof of Lemma EC.4. Fix a nonanticipative IA-feasible policy π and couple it to $\mathcal{DA}(\pi)$ through the same realization of π's internal randomness. Along a joint realization of demands and internal randomness, write $d_{i,t}$, $\tilde{q}_i^{(t-1)}$, $z_{i,t}$, and $x_{i,t}$ for the corresponding product-i demand, uncommitted residual inventory, bag assignment, and regular display. Define $Z_{i,t} := \sum_{u=1}^{t} z_{i,u}$. The shadow's physical inventory of product i immediately before period t is $\tilde{q}_i^{(t-1)} + Z_{i,t-1}$. After period-$t$ sales, it is

$$\tilde{q}_i^{(t-1)} + Z_{i,t-1} - \min\{d_{i,t}, x_{i,t}\} = \tilde{q}_i^{(t)} + Z_{i,t}, \qquad i \in [N].$$

Thus IA feasibility makes every shadow display feasible. At the horizon, each product's physical inventory is its uncommitted residual inventory plus its cumulative deferred assignments, so the latter can be fulfilled without composition constraints. After fulfilling the cumulative deferred assignments, the shadow retains exactly $\tilde{q}_i^{(T)}$ units of product i, which is also the inventory discarded by the IA policy at the horizon.

The shadow observes exactly the same information as π before each period's demand realization. Therefore the nonanticipativity of π implies the nonanticipativity of every shadow decision. Finally, the shadow reproduces the same regular sales, bag revenue, and waste at the horizon as π on every joint realization. Hence the two policies have identical pathwise profit. □

EC.8. Dual Derivations for the Static and Branch-Free Linear Programs

In this section, we derive the duals of static LP and the uniform-gain LP. We suppress the constant baseline profit $\mathbb{E}[\textsc{Profit}[\mathrm{RO}; I]]$ and add it back at the end. Both LPs fit the same generic form

$$\max_{u} \sum_{(i,k)\in\mathcal{U}} c_{i,k}\, u_{i,k} \tag{EC.29}$$

$$\text{s.t.} \sum_{(i,k)\in\mathcal{U}} (\mathbf{a}_i)_\ell\, u_{i,k} \le (\mathbf{b})_\ell \qquad \forall \ell \in \{0\} \cup [L],$$

$$u_{i,k} \le 1 \qquad \forall (i,k) \in \mathcal{U},$$

$$u_{i,k} \ge 0 \qquad \forall (i,k) \in \mathcal{U} \tag{EC.30}$$

with $(\mathbf{a}_i)_0 = 1$, $(\mathbf{a}_i)_\ell = a_{i,\ell}$ for $\ell \in [L]$, $(\mathbf{b})_0 = TY$, and $(\mathbf{b})_\ell = b_\ell$ for $\ell \in [L]$.

Dual derivation. Introduce dual variables

$$\lambda_\ell \geq 0 \quad \forall \ell \in \{0\} \cup [L], \qquad \mu_{i,k} \geq 0 \quad \forall (i,k) \in \mathcal{U}.$$

The dual of (EC.29)–(EC.30) is

$$\begin{aligned} \min \quad & \sum_{\ell \in \{0\}\cup[L]} (\mathbf{b})_\ell \lambda_\ell + \sum_{(i,k)\in\mathcal{U}} \mu_{i,k} \\ \text{s.t.} \quad & \mu_{i,k} + \sum_{\ell \in \{0\}\cup[L]} (\mathbf{a}_i)_\ell \lambda_\ell \geq c_{i,k} \qquad \forall (i,k) \in \mathcal{U}, \\ & \lambda_\ell \geq 0 \qquad \forall \ell \in \{0\} \cup [L], \\ & \mu_{i,k} \geq 0 \qquad \forall (i,k) \in \mathcal{U}. \end{aligned}$$

Reduced dual. Fix $\lambda \geq 0$. The smallest feasible $\mu_{i,k}$ is

$$\mu_{i,k} = \Big(c_{i,k} - \sum_{\ell \in \{0\}\cup[L]} (\mathbf{a}_i)_\ell \lambda_\ell\Big)^+.$$

Substituting into the objective yields

$$\min_{\lambda \in [0,\infty)^{L+1}} \left\{ \sum_{\ell \in \{0\}\cup[L]} (\mathbf{b})_\ell \lambda_\ell + \sum_{(i,k)\in\mathcal{U}} \Big(c_{i,k} - \sum_{\ell \in \{0\}\cup[L]} (\mathbf{a}_i)_\ell \lambda_\ell\Big)^+ \right\}. \tag{EC.31}$$

LP specializations. For both LPs, the zero vector is feasible and the box constraints make the feasible region bounded. Hence, each primal has a finite optimal value, and LP strong duality equates it with the value of the corresponding reduced dual. Setting $c_{i,k} = s - v_{i,k}$ in (EC.31) gives the dual of the static LP:

$$\mathrm{LP}_{\mathrm{st}} = \mathbb{E}[\textsc{Profit}[\mathrm{RO}; I]] + \min_{\lambda \in [0,\infty)^{L+1}} \left\{ \sum_{\ell \in \{0\}\cup[L]} (\mathbf{b})_\ell \lambda_\ell + \sum_{(i,k)\in\mathcal{U}} \Big(s - v_{i,k} - \sum_{\ell \in \{0\}\cup[L]} (\mathbf{a}_i)_\ell \lambda_\ell\Big)^+ \right\}.$$

Setting $c_{i,k} = s + w$ in (EC.31) gives the dual of the uniform-gain LP:

$$\mathrm{LP}_{\mathrm{ug}} = \mathbb{E}[\textsc{Profit}[\mathrm{RO}; I]] + \min_{\lambda \in [0,\infty)^{L+1}} \left\{ \sum_{\ell \in \{0\}\cup[L]} (\mathbf{b})_\ell \lambda_\ell + \sum_{(i,k)\in\mathcal{U}} \Big(s + w - \sum_{\ell \in \{0\}\cup[L]} (\mathbf{a}_i)_\ell \lambda_\ell\Big)^+ \right\}.$$

EC.9. Heuristics and Experimental Benchmarks

This section describes the two heuristic policies and the information-relaxed benchmark used for empirical comparison in Section 6. The two heuristics are restricted versions of the LP-based reservation policy: a uniform-share reservation rule and a certainty-equivalent value rule.

EC.9.1. Uniform-share policy

The uniform-share policy uses the same bag-channel proportion across products:

$$\begin{aligned}\max_{h\in[0,1],\,u\in[0,1]^{\mathcal{U}}} \quad & \sum\nolimits_{(i,k)\in\mathcal{U}}(s-v_{i,k})u_{i,k}\\ \text{s.t.} \quad & \sum\nolimits_{(i,k)\in\mathcal{U}} u_{i,k}\le TY,\\ & \sum\nolimits_{i\in[N]} a_{i,\ell}\sum\nolimits_{k=1}^{q_i} u_{i,k}\le b_\ell, \qquad \ell\in[L],\\ & \sum\nolimits_{k\in[q_i]} u_{i,k}=h\,q_i, \qquad i\in[N]. \end{aligned} \tag{EC.32}$$

The uniformity constraint (EC.32) means that each product routes the same fraction h of its initial inventory to the surprise-bag channel. This is the only difference from the LP-based reservation policy. We then apply the same rounding step and online execution as in the LP-based reservation policy.

EC.9.2. Certainty-equivalent policy

The certainty-equivalent policy replaces the unit-level regular values $v_{i,k}$ with certainty-equivalent regular values computed from mean demand, for any $i\in[N]$ and $k\in[q_i]$. Let

$$\bar{R}_i \triangleq \min\Big\{q_i,\ \sum\nolimits_{t=1}^{T}\min\big\{\mathbb{E}[d_{i,t}]\,,\,X_i\big\}\Big\} = \big\lfloor \bar{R}_i\big\rfloor + \bar{f}_i,$$

where $\bar{f}_i \triangleq \bar{R}_i - \big\lfloor \bar{R}_i\big\rfloor \in [0,1)$. The corresponding certainty-equivalent regular value is

$$\bar{v}_{i,k} = \begin{cases} r_i, & k\le \big\lfloor \bar{R}_i\big\rfloor,\\ \bar{f}_i\,r_i-(1-\bar{f}_i)w, & k=\big\lfloor \bar{R}_i\big\rfloor+1,\\ -w, & k\ge \big\lfloor \bar{R}_i\big\rfloor+2. \end{cases}$$

We then solve the same static LP as in the LP-based reservation policy, with $\bar{v}_{i,k}$ in place of $v_{i,k}$, and apply the same rounding step and online execution as in the LP-based reservation policy.

EC.9.3. Ex-ante benchmark

The ex-ante benchmark is an information relaxation used only for evaluation. Recall that $p_{i,k}$ is the probability that unit (i,k) would sell under RO. For each unit, let $u^+_{i,k}$ and $u^-_{i,k}$ denote its conditional probabilities of being allocated to a surprise bag, given that it would and would not sell under RO, respectively. The value of a variable on a zero-probability branch is immaterial. We define

$$\begin{aligned}\mathrm{LP}_{\mathrm{ea}} \triangleq \mathbb{E}[\textsc{Profit}[\mathrm{RO};I]] + \max_{u^+,u^-\in[0,1]^{\mathcal{U}}} \quad & \sum\nolimits_{(i,k)\in\mathcal{U}}\Big[p_{i,k}(s-r_i)u^+_{i,k}+(1-p_{i,k})(s+w)u^-_{i,k}\Big]\\ \text{s.t.} \quad & \sum\nolimits_{(i,k)\in\mathcal{U}}\Big(p_{i,k}u^+_{i,k}+(1-p_{i,k})u^-_{i,k}\Big)\le TY,\\ & \sum\nolimits_{(i,k)\in\mathcal{U}} a_{i,\ell}\Big(p_{i,k}u^+_{i,k}+(1-p_{i,k})u^-_{i,k}\Big)\le b_\ell, \qquad \ell\in[L]. \end{aligned} \tag{EC.33}$$

Here, $p_{i,k}u^+_{i,k}+(1-p_{i,k})u^-_{i,k}$ is the unconditional probability that unit (i,k) is allocated to a surprise bag.

For any feasible policy, set $u^+_{i,k}$ and $u^-_{i,k}$ to its corresponding conditional allocation probabilities. The law of total expectation and the policy's feasibility imply that these variables satisfy the constraints above. Relative to RO, allocating a unit that would sell regularly increases profit by at most $s - r_i$, whereas allocating a unit that would remain unsold increases profit by $s + w$. The objective at the induced solution therefore upper-bounds the policy's expected profit. Thus, $\mathrm{LP}_{\mathrm{ea}}$ upper-bounds the expected profit of every feasible policy; it is a benchmark value rather than an implementable policy.