Discussion Paper Series – CRC TR 224

Discussion Paper No. 083
Project B 02

Revenue Management without Commitment: Dynamic Pricing and Periodic Flash Sales

Francesc Dilmé*
Fei Li**

April 2019

*University of Bonn, fdilme@uni-bonn.de
**University of North Carolina, Chapel Hill, lifei@email.unc.edu

Funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 is gratefully acknowledged.

Revenue Management without Commitment: Dynamic Pricing and Periodic Flash Sales*

Francesc Dilmé† Fei Li‡

April 11, 2019

Abstract

A seller has a fixed number of goods to sell by a deadline. At each time, he posts a regular price and decides whether to hold a flash sale. Over time, buyers privately enter the market and strategically time their purchases. If a buyer does not purchase when she arrives, she can pay an attention cost to recheck the regular price afterwards, or she can wait for future flash sales where she may obtain a good at a discounted price. In the unique Markov perfect equilibrium, the seller sporadically holds flash sales to lower the stock of goods. A flash sale increases the willingness to pay of future buyers, but decreases the willingness to pay of buyers who arrive early in the game. When it is very likely that a buyer will obtain a good in a flash sale, the seller holds a "big" initial flash sale for all but one unit of the good.

Keywords: revenue management, commitment power, dynamic pricing, flash sales, inattention frictions.

JEL Classification Codes: D82, D83

*We are grateful to George Mailath and Mallesh Pai for their guidance and advice. We also thank editor Botond Koszegi, three anonymous referees, Aislinn Bohren, Heski Bar-Isaac, Gary Biglaiser, Simon Board, Luis Cabral, Eduardo Faingold, Hanming Fang, George Georgiadis, Johannes Horner, Daniel Krähmer, Kyungmin Kim, Stephan Lauermann, Anqi Li, Qingmin Liu (discussant), Steven Matthews, Benny Moldovanu, Guido Menzio, Peter Norman, Alessandro Pavan, Andrew Postlewaite, Maher Said (discussant), Yuliy Sannikov, Philipp Strack, Tymon Tatur, Can Tian, Rakesh Vohra, Jidong Zhou, and participants at numerous conferences and seminars for their valuable comments. Any remaining errors are ours. Dilmé thanks the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) for research support through grant CRC TR 224.

†University of Bonn, fdilme@uni-bonn.de

‡University of North Carolina, Chapel Hill, lifei@email.unc.edu

1 Introduction

Some markets share the following characteristics: (1) goods are perishable, (2) the initial number of goods for sale is fixed in advance, and (3) consumers enter the market over time and choose the timing of their purchases. Examples of such markets include airline seats, hotel rooms, and sports tickets. The revenue management literature studies the optimal pricing in these markets under the assumption that the seller has perfect commitment power. By contrast, this paper studies the case where the seller cannot commit.[1]

In our model, a monopolist seller has a fixed number of identical goods to sell before a deadline. At each time, the seller posts a price and also decides whether to offer some goods in an advertised flash sale featuring a low price. There are two kinds of consumers: arriving (high-valuation) buyers and a stationary set of (low-valuation) shoppers. Arriving buyers are forward-looking and have unit demand, and their arrival is private. When a buyer arrives, she observes the current price and the remaining inventory and decides whether to buy a good. If she decides not to make the purchase, she becomes an *accumulated buyer* and faces *inattention frictions*: she can pay costly attention to recheck the price and remaining stock at any time. Flash sales attract the attention of both accumulated buyers and shoppers.[2] In a flash sale, the probability that an accumulated buyer obtains a unit depends on the number of units offered by the seller.

We characterize the unique Markov perfect equilibrium where the behavior of the players depends only on commonly available information: the calendar time and the number of remaining goods of the seller. In equilibrium, due to the lack of commitment, the seller holds a flash sale at the deadline to clear the remaining stock.[3] Prior to the deadline, forward-looking buyers anticipate that flash sales may be offered in the future, and so they expect to obtain a positive surplus

[1] In the airline and hotel industries, sellers often employ revenue management algorithms, a practice typically viewed as a commitment device used by sellers. However, revenue managers frequently adjust prices based on their information and personal experience, instead of strictly following the pricing strategy suggested by the algorithm. In the article *Confessions of an Airline Revenue Manager*, George Hobica said,"The computer adjusts fares all the way up until the departure time, but as a revenue manager, I can go in and adjust things based on information that the computer may not know. For example, are there specific events taking place at a destination? Are there certain conditions at the departure airport that will allow more than the desired amount of seats to go empty such as weather?" See http://www.foxnews.com/travel/2011/12/08/confessions-airline-revenue-manager.html

[2] Flash sales are common in many industries. For example, in the airline industry, flash sales (limited time offers with sufficiently low prices) are common. See https://www.skyscanner.net/news/tips/how-to-find-flash-sale-flights-before-theyre-gone/. Casual observation suggests that these offers are advertised to "window shoppers" via deal-alert email/text messages from agencies like Kayak, Orbitz, or Google Flights or online real-time bidding sponsored advertisement on social media such as YouTube and Facebook.

[3] Last-minute clearance sales are optimal in many dynamic pricing settings. See, e.g., Nocke and Peitz (2007). In practice, the last-minute deal strategy is commonly used in many industries. See *The Wall Street Journal*, March 15, 2002, "Airlines now offer 'last-minute' fare bargains weeks before flights," by Kortney Stringer.

by delaying their purchases. However, waiting for flash sales is risky: goods may be purchased by others. To avoid such a risk, a buyer is willing to purchase a good immediately at a price higher than the flash-sale price. The seller posts the endogenous reservation price of arriving buyers, ensuring that they purchase a good upon arrival. As the deadline approaches, a buyer's willingness to pay decreases because the arrival of new buyers becomes less likely. Also, at any time, the smaller the seller's inventory, the higher the risk a buyer faces by waiting for flash sales, and thus the higher the price she is willing to pay to avoid the risk. Consequently, the equilibrium price decreases continuously over time until a transaction occurs, the inventory is reduced, and the price jumps up.

A novel feature of our equilibrium is that the seller periodically holds flash sales before the deadline. In doing so, the seller lowers his inventory and creates scarcity, which allows him to charge a high price in the future. The optimal timing of the flash sales trades off between future transaction prices and quantities: by reducing the current inventory size, the seller can extract more rent from future buyers, but he increases the probability that the remaining stock is insufficient to sell to all of them. In equilibrium, it is unlikely that a large number of buyers will arrive before the deadline when it is near, and so the seller has an incentive to keep a small inventory. The equilibrium specifies a sequence of "cutoff times" when the seller finds it optimal to hold a flash sale if the remaining stock is high.

Our first contribution is to construct a tractable setting that makes it possible to study dynamic price discrimination in such an environment. The main modeling innovation is the presence of inattention frictions on the buyers' side. Our modeling of inattention frictions helps us to overcome two significant challenges in modeling revenue management with hidden arrivals: dealing with both the large set of deviating strategies and the possession of private information by the different agents in the model.[4] In general, when arrivals are private, forward-looking agents need to form beliefs about both the number of accumulated buyers and their private information. Accordingly, the seller can use prices to manipulate buyers' beliefs and their willingness to pay. This not only makes characterizing equilibrium behavior difficult, but may also lead to unreasonable (on- or off-path) equilibrium behavior supported through implausible coordination between the agents' beliefs about it. Our modeling of inattention frictions limits the seller's ability to manipulate buyers' beliefs, as inattentive buyers do not closely track the price. Also, flash sales allow the seller to sell units fast at low prices, but since they attract the attention of accumulated buy-

[4]We perceive inattention frictions as a natural way to capture relevant aspects of consumer behavior. For example, due to the physical and psychological costs, when a consumer is waiting for a discount flight ticket, she will only occasionally visit the airline's website instead of monitoring it 24-7.

ers, they restrict the ability of the seller to use the inattention frictions to fully price discriminate between buyers and shoppers.

Our second contribution is to identify a new channel of inefficiency in dynamic monopoly pricing: the seller allocates some of the goods early in the game through flash sales despite the possibility that additional buyers with high willingness to pay may arrive in the future.

The third contribution of the paper is to provide a new rationale for price fluctuation and flash sales. Holding flash sales creates scarcity for future consumers and allows the seller to extract more rent from them. This complements the classic view of price fluctuation (discussed in the literature review) as stemming from either changes of the inventory or the seller's screening of buyers with different valuations and arrival times.[5] One of the testable implications distinguishing our model from the literature is that price jumps do not have to take place either after transactions (as in most revenue management models) or when a sufficient number of low-value consumers are accumulated (as in the durable goods literature). They can also occur when the realized demand is low and the deadline is close.

The aforementioned results rely on the following building blocks of our model: the existence of a deadline, buyers arriving stochastically and being forward-looking, the finiteness of the inventory, and the lack of commitment by the seller. First, as the buyers' arrival is stochastic, there are histories where the stock is high and the expected future demand from arriving buyers is low due to the proximity of the deadline. The presence of the deadline is crucial for the seller to be willing to hold a flash sale at such histories. Alternatively, if the arrival times of the buyers were deterministic, the seller would hold a flash sale (if at all) only at the beginning of the game, since a late flash sale would reduce the willingness to pay of buyers who arrive early. Second, the seller's incentive to hold flash sales also relies on buyers being forward-looking. If, on the contrary, buyers were myopic, their willingness to pay would be independent of the seller's inventory. Third, a flash sale can credibly reduce future supply only if the seller is unable to produce additional goods over time, which is ensured by the assumption of the inventory being finite. Finally, the lack of commitment by the seller contributes to shaping the price dynamics. In the equilibrium, the seller is unable to extract the full surplus from the buyers, as they believe that the seller will hold low-price flash sales in the future. If the seller had commitment power, he could commit to not holding flash sales at all. He benefits from doing so if, for example, the flash-sale price is sufficiently small.

Finally, we investigate the effect of demand-side competition on price dynamics. A buyer's

[5]For example, there is extensive evidence of price fluctuation in the airline market. McAfee and Te Velde (2006) find that the fluctuation of airfares is too high to be explained by the standard monopoly pricing models. Escobari (2012) and Williams (2018) find that airline prices decrease conditional on inventory size and increase after transactions.

willingness to pay depends on the probability that she will get a good in a future flash-sale offer if she rejects the price offer at her arrival time. This probability quantifies the competition between buyers and shoppers during flash sales. We first consider the case where, whenever there is one accumulated buyer and the seller holds a flash sale, the buyer is very likely to acquire a good. In this case, the buyers' willingness to pay is very low shortly before a flash sale, and so the seller has an incentive to hold a flash sale earlier to increase the price. We show that when the probability that a buyer obtains a good in a flash sale is high, the seller holds a flash sale for all but one unit at the beginning of the game, and tries to sell the remaining unit at the regular price. Then we let accumulated buyers become increasingly unlikely to obtain a good in a flash sale. At the limit where a buyer cannot obtain a good in a flash sale, the seller's commitment problem disappears: he charges a high price, leaves no rent for buyers, and holds a flash sale to sell the remaining units at the deadline.

Related Literature. A large body of the revenue management literature examines markets with sellers who sells a finite number of goods before a deadline to myopic buyers who arrive over time.[6] The common finding in these models is that prices fluctuate over time, decreasing before sales take place and sharply increasing afterwards. In a recent paper, Board and Skrzypacz (2016) extend the analysis to forward-looking buyers and characterize the revenue-maximizing mechanism.[7] The optimal mechanism trades off between time discounting and incoming demand with higher willingness to pay. In the continuous-time limit, the optimal mechanism is implemented via a price-posting mechanism with a last-minute auction. Our findings in the absence of commitment have three main distinguishing features. First, while in the commitment case the main inefficiency arises because of monopolistic withholding, in our model it stems from the early flash sales. Second, price fluctuations in our model arise even in the absence of discounting and are caused by the inability of the seller not to lower future prices. Third, Board and Skrzypacz argue that the seller's profit is higher when buyers are long-lived, as the commitment power gives the seller the option to postpone transactions and wait for the resolution of arrival uncertainty. On the contrary, in our model where the seller has no commitment power, the seller's profit is lower when buyers are long-lived due to the equilibrium strategic interaction between them (see Proposition 6). Finally, in Board and Skrzypacz, the seller has no incentive to hide the inventory or manipulate the buyers' belief about it. By contrast, manipulating the inventory is crucial for the seller in our model, as it is the only credible way he has to influence the buyer's willingness to pay.

[6] See, e.g., the seminal paper of Gallego and Van Ryzin (1994), and Talluri and Van Ryzin (2006) for a textbook treatment. Gershkov and Moldovanu (2009) extend the benchmark model to address the case of heterogeneous objects.

[7] See also Mierendorff (2016), Pai and Vohra (2013), and Gershkov et al. (2017).

In Hörner and Samuelson (2011) and Chen (2012), the seller has no commitment power and sells goods to a fixed set of forward-looking buyers. In a setting without randomly arriving buyers, they show that the seller either replicates a Dutch auction or posts an unacceptable price until the deadline and then charges a static monopoly price. Therefore, in their models, the seller never uses low prices to inefficiently sell to buyers with low valuations to increase the competition among buyers with high valuations. By contrast, we show that maintaining the assumption that buyers arrive stochastically, as do most revenue management models, generates both fluctuating price dynamics and the possibility of flash sales. Deb and Said (2015) allow the seller to manage demand–as opposed to supply in our model–so that he can manage competition among consumers inter-temporally. They consider a two-period dynamic screening problem where a seller without capacity constraints sells goods to buyers who arrive in each period. They show that the seller may deliberately postpone the transaction with some first-period buyers to alter demand in the second period, thereby making the threat of charging a high second-period price credible.

More broadly, our paper is related to the durable-goods pricing literature. First, a set of papers introduce arrival of consumers into a canonical durable good pricing model à la Gul et al. (1986). Conlisk et al. (1984) and Sobel (1991) study pricing dynamics with the arrival of new buyers per period. In their model, the price fluctuates due to the trade-off between rent extraction from high-valuation buyers and the revenue from the accumulated low-valuation buyers.[8] Öry (2017) examines a model where it is costly for the seller to lure "window shoppers" back. Similarly to our model, her model exhibits discrete price drops at the times when the seller sells to low-valuation buyers. Although the aforementioned papers also generate price fluctuations as in our model, the key mechanism is different. In the durable goods literature, the low-value buyers are accumulated until a critical time at which the seller finds it optimal to reap the rent from them. Instead, in our model, the seller holds flash sales to increase the scarcity of the goods, consequently increasing the willingness to pay of future buyers after low realizations of the demand. Therefore, our model is better suited to explain price fluctuation in environments where sellers have binding inventory constraints and goods are perishable. Second, the role of capacity constraint has also been studied in durable-goods pricing contexts. McAfee and Wiseman (2008) introduce endogenous choice of capacity in each period and allow the production cost of capacity to be vanishingly small. They show the Coase conjecture fails in the presence of capacity constraints rather than to rationalize the price fluctuation. McAfee and Vincent (1997) assume that the seller has one good and posts the

[8] Garrett (2016) considers a model where buyers arrive privately and their valuations change stochastically over time. He shows that similar price dynamics take place. Board (2008) discusses the case where incoming demand varies over time and examines how the seller uses time to discriminate between different generations. Also see Inderst (2008) and Fuchs and Skrzypacz (2013) for studies on arrivals in bilateral bargaining with asymmetric information.

reserve price of an auction in each period. When there is a "gap" between the seller's reservation value and minimum valuation of the buyer, they show that the seller's expected revenue converges to those of a static auction with no reserve price when the time between auctions goes to zero. Liu et al. (2018) study the "no gap" case and examine both stationary and non-stationary equilibria. They show that with multiple buyers, an immediate sale by efficient auction is optimal in the limit as the period length vanishes.

There is a large industrial organization literature on the roles of capacity constraint and of demand uncertainty in pricing. For example, Gale and Holmes (1993) show that it is optimal for a monopoly airline to use advance-purchase discounts to divert consumers from a peak-time flight to an off-peak-time flight. Dana (1998) discusses the optimality of advance-purchase discounts in a perfectly competitive market à la Eden (1990). Also noteworthy are the studies on the effect of costly capacity choice on price competition, e.g., Kreps and Scheinkman (1983), Dana (1999), Anton et al. (2014) and Dana and Williams (2018). In this literature, it is typical to investigate a static or two-period stylized model, and the focus is not on fully analyzing the price dynamics as in our model.

Finally, the insight that sellers may benefit from reducing inventories or endowments has been discussed in static competitive equilibrium frameworks introduced by Aumann and Peleg (1974) and Gale (1974). Our goal is to investigate how such an insight shapes the equilibrium outcome in a revenue management environment, where dynamic considerations and random demand play an important role in pricing. In particular, we focus on the *timing* of the flash sales.

The rest of this paper is organized as follows. In Section 2, we present the model setting and define the solution concept we are going to use. In Section 3, we analyze equilibria. Section 4 discusses some modeling choices and possible extensions of the baseline model. Appendix A includes all the proofs.

2 Model

We consider a dynamic pricing game between a single seller who has $K \in \mathbb{N}$ identical and indivisible goods for sale and many buyers. Time is continuous: $t \in [0,1]$.

Seller. At every time t, the seller makes a *regular offer* which specifies a price $P_t \in \mathbb{R}$. After offering the regular price, he can hold a *flash sale*, discussed in more detail below. The seller values goods at zero.

Buyers. At time 0, there are no buyers. Over time, buyers arrive privately at a rate $\lambda > 0$. Each buyer has a unit demand and values the good at $v_H > 0$. A buyer leaves the game after purchasing a good. When a buyer arrives, she observes the current regular offer P_t and the seller's current *inventory* (or stock) K_t, which equals the number of units that have not been sold in $[0, t)$. She either accepts the regular offer, or she rejects it and becomes an *accumulated buyer*. An accumulated buyer faces *inattention frictions*: at each moment t, she does not observe the seller's regular offer and inventory unless she incurs a cost $c > 0$ for paying attention. In sum, a buyer is *attentive* at her arrival time and at the times when she pays attention (or *rechecks* the regular price).[9]

If there are $N_t \geq 1$ attentive buyers at time t, they simultaneously and independently decide whether to purchase. Nature randomly orders the buyers who decide to purchase a good and allocates the goods to the first K_t of them (or to all of them, if fewer than K_t buyers are willing to buy a good).[10] These are called *regular sales*.

Flash sales. At each time t, after the regular sales take place, the seller decides the number of units to be offered in a *flash sale*. A flash sale differs from a regular offer in three aspects: (i) it is held at an exogenous price $v_L \in (0, v_H)$, (ii) without paying an attention cost, accumulated buyers observe it with a positive probability, and (iii) it also attracts demand from a broader market. One can interpret such additional demand as a pool of shoppers with a willingness to pay equal to v_L.

If the seller chooses to hold a flash sale, he selects the number of units to be offered in it. The game proceeds similarly to the regular price offer, but with the possibility that some goods will be purchased by shoppers. First, nature uniformly randomly orders the accumulated buyers (if any). If there is no buyer, all goods are assigned to shoppers. If there is at least one buyer, with probability $\beta \in (0, 1)$ the first good is assigned to the first buyer, and with probability $1 - \beta$ to a shopper. If more than one unit is offered in the flash sale, the second unit is assigned analogously: if no buyer remains, it is assigned to a shopper, while if there is at least one buyer, the unit is assigned to the first of the remaining buyers with probability β, and with probability $1-\beta$ to a shopper. This process is iterated until none of the units offered in the flash sale is left. The parameter β measures how "attentive" buyers are with respect to shoppers. It also captures the seller's inability to exclude accumulated buyers from having access to flash sales, making the seller's price discrimination problem non-trivial. The seller does not know whether the goods are assigned to buyers or shoppers, and a buyer is aware of the flash sale only if a good is assigned to her.

Information. The seller observes his own previous actions (the previous regular offers and flash sales) and the previous regular sales. He does not observe the arrival of buyers, their attention

[9] The role of the assumption of inattention frictions is discussed in Section 2.1.

[10] Our results are unchanged if the seller is allowed to control the capacity for each regular price.

times, or the type of consumers who obtained goods in each flash sale. Buyers observe the regular price and the inventory at each of their attention times. They do not observe the arrival of other buyers or whether flash sales take place. A buyer observes a flash-sale offer only if a unit is assigned to her.

Heuristic timing and payoffs. Although the game is in continuous time, it is useful to consider a heuristic internal timing within each instant t as follows: (i) the seller chooses regular price P_t, (ii) nature decides whether a buyer arrives at time t, and accumulated buyers make their attention choice, (iii) attentive buyers observe (K_t, P_t) and decide whether to accept the current regular offer, (iv) goods are allocated according to the regular-sales procedure described above, (v) knowing how many goods have been sold at the regular price, the seller decides how many units to offer in a flash sale, and goods are allocated according to the flash-sale procedure described above.

The seller is risk-neutral and does not discount the future. His payoff equals the summation of all transaction prices. A buyer's payoff is given by

$$\begin{cases} v_H - p - nc & \text{if she purchases the good at a price } p\,, \\ -nc & \text{if she fails to get any good}\,, \end{cases}$$

if she pays costly attention for $n \in \mathbb{Z}_+$ times.

Histories and strategies. To avoid technical issues associated with continuous-time games (see Bergin and MacLeod, 1993), we follow Murto and Välimäki (2013) and Akcigit and Liu (2015) by endowing the players' strategies with multiple private stages. A player's private stage begins when he/she obtains new information regarding other players' actions. One can then interpret the strategy of each player as deciding, at the beginning of each of his/her private stages, his/her action plan until the stage ends. This specification allows each player to immediately react to other players' actions. The next four paragraphs provide the formal definitions of the history and strategy sets of each player. They enable us to define Markov strategies and Markov equilibria, which are the basis of our later analysis.

A *seller's history* at time t summarizes information about previous regular sales up to t and the remaining stock $K_t \geq 0$. If, before time t, regular sales have taken place for $I \in \{1, ..., K\}$ times, the seller's history consists of a strictly increasing, finite sequence of times $(\tau_1, \tau_2, ..., \tau_I < t)$ when the previous regular sales took place and of the volume of sales. Hence, the seller's history up to time t is denoted by $h_t^s = ((t, K_t), (\tau_i, r_i)_{i=0}^I)$, where $r_i \geq 1$ represents the volume of regular sales at time τ_i, for $i = 1, ..., I$, and where we set $(\tau_0, r_0) = (0, 0)$ for convenience. We require h_t^s to satisfy

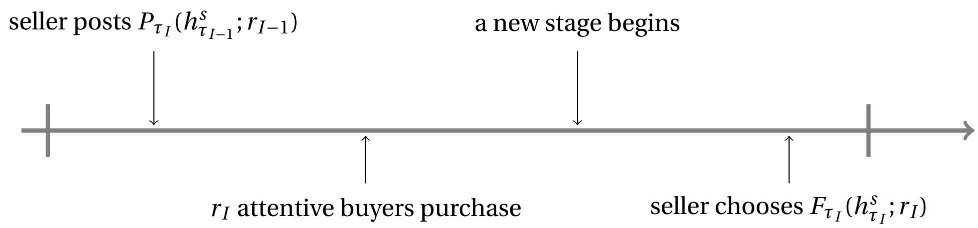

Figure 1: A heuristic order of moves at τ_I when a regular sale happens and a new stage begins, where $\tau_{I-1} < \tau_I$ denotes the last time a regular transaction occurred.

$K_t \le K - \sum_{i=1}^{I} r_i$. If, instead, there have been no sales up to t the history at time t, we set $h^s_t = ((t, K), (0,0))$.

A *seller's (pure) strategy* is defined in stages as follows. At time 0, the seller begins with private stage 0 at history $h^s_0 = ((0,K),(0,0))$. The seller's strategy in stage 0 specifies a regular pricing and flash-sale decision for each $t \ge 0$. Specifically, it specifies for each $t \ge 0$ (i) $P_t(h^s_0) \in \mathbb{R}_+$, the regular price at time t conditional on no regular sale having occurred in $[0,t)$, and (ii) $F_t(h^s_0) \in \{0,\dots,K\}$, the cumulative number of flash sales in $[0,t]$ conditional on no regular sale having occurred in $[0,t]$. Whenever a regular sale occurs at time $\tau_I \in [0,1]$, the seller enters the subsequent private stage. Using $h^s_{\tau_I} = ((\tau_I, K_{\tau_I}), (\tau_i, r_i)_{i=0}^{I-1})$ to denote the seller's history at the beginning of the stage that starts at time τ_I and r_I to denote the number of regular sales at τ_I, the seller's strategy specifies (i) the regular price $P_t(h^s_{\tau_I}; r_I) \in \mathbb{R}_+$ for each $t \in (\tau_I, 1]$ and (ii) the cumulative number of flash sales $F_t(h^s_{\tau_I}; r_I) \in \{0,\dots,K_{\tau_I} - r_I\}$ in $[\tau_I, t]$.[11] To guarantee that the times when flash sales occur are well defined, both $F_t(h^s_0)$ and $F_t(h^s_{\tau_I}; r_I)$ are assumed to be a right-continuous and non-decreasing function of t in their domain. Figure 1 illustrates a typical sequence of moves when a new private stage of the seller is initiated by buyers' purchase at time τ_I.

A *buyer's history* is defined only when she is attentive. At time t, an attentive buyer's history consists of the previous and current attention times, $(\tau^b_1 = b, \tau^b_2, \dots, \tau^b_J = t)$, and the prices and the inventories at these times $(P_j, K_j)_{j=1}^J$, where b is the buyer's arrival time. Denote it by $h^b_\tau \equiv (\tau^b_j, P_j, K_j)_{j=1}^J$. A buyer begins her first stage at her arrival time, and she enters the next private stage whenever she rechecks the regular price. At the beginning of her J-th stage, given her history h^b_t, the *buyer's (pure) strategy* specifies a pair: the first component $D(h^b_t) \in \{0,1\}$ indicates whether the buyer should accept the regular offer at time t, and the second component represents the

[11]Notice that the domain of $F_t(h^s_{\tau_I}; r_I)$ includes τ_I, that is, the seller can immediately determine the number of flash sales at the moment a regular sale occurs when a new stage begins. Therefore, the seller can react to a buyer's purchase at time t without a lag. Also, if $I > 0$, $P_t(h^s_{\tau_I}; r_I)$ is defined for $t > \tau_I$; while if $I = 0$, the domain of t in $P_t(h^s_0)$ includes 0 so that the regular price at time 0 is specified. This asymmetric treatment is due to the fact that private stage 0 starts before the seller's regular pricing decision at $t = 0$.

next attention time $A(h_t^b) \in (t,1] \cup \{+\infty\}$ conditional on not obtaining a good at time t, where $+\infty$ means that she will not check the price in the future. It is convenient to assume that any strategy of a buyer generates a finite number of rechecking times, as it ensures that (i) the history of a buyer is finite dimensional, and (ii) the payoff of a buyer is finite independently of the strategy she uses.[12]

Solution concept. We focus on equilibria where the strategies of the players depend on the calendar time and the inventory size.[13] With some abuse of language, we say that the seller's strategy is *Markov* if there is a regular price $p_k(t) \in \mathbb{R}_+$ and a flash-sale policy $f_k(t) \in \{0,...,k\}$ for each $t \in [0,1]$ and $k \in \{1,...,K\}$ satisfying the following property. For each stage beginning with a seller's history $h^s_{\tau_I} = ((\tau_I, K_{\tau_I}), (\tau_i, r_i)_{i=0}^{I-1})$ and an r_I-unit regular sale,

- $F_{\tau_I}(h^s_{\tau_I}; r_I) = f_{K_{\tau_I}-r_I}(\tau_I)$, and
- $P_t(h^s_{\tau_I}; r_I) = p_k(t)$ and $F_t(h^s_{\tau_I}; r_I) - F_{t^-}(h^s_{\tau_I}; r_I) = f_k(t)$ for every $t \in (\tau_I, 1]$, where $k = K_{\tau_I} - r_I - F_{t^-}(h^s_{\tau_I}; r_I)$.

Hence, in a Markov strategy of the seller, $p_k(t)$ denotes the regular price at time t when the seller has k units, and $f_{k-r}(t)$ denotes the number of units offered in the time-t flash sale if $r \in \{0,...,k\}$ units were sold in regular sales at time t. Similarly, a strategy of a buyer is *Markov* if, for any buyer's history h_t^b, $D(h_t^b)$ and $A(h_t^b)$ depend only on (t, K_t) and P_t. We say that a strategy profile is a *Markov perfect equilibrium* if it is a perfect Bayesian equilibrium (in pure strategies) and the players' strategies are Markov.

In our game, players have a significant amount of private information regarding previous prices and accumulated buyers even in a Markov equilibrium. As buyers' decisions depend on not only the Markov state variable but also the regular price, the seller may use the price to signal his (endogenous) private information to manipulate subsequent buyers' purchasing and attention decisions off the equilibrium path. To isolate our main focus on dynamic price discrimination from technical complications driven by signaling considerations and the evolution of beliefs about the number of accumulated buyers, we select Markov perfect equilibria where (i) there is no accumulated buyer on the path, and (ii) after *any* buyer's history, a buyer believes that there is no other

[12] Formally, a buyer's strategy (A,D) is feasible if the following holds. Fix any path $(P_t, K_t)_{t=0}^1$ and arrival time b. Define $\tau_1^b = b$ and $h_b^b = (b, P_b, K_b)$. Recursively define $\tau_j = A(h^b_{\tau^b_{j-1}})$ and $h^b_{\tau^b_j} = (h^b_{\tau^b_{j-1}}, (\tau^b_j, P_{\tau^b_j}, K_{\tau^b_j}))$ for all $j > 1$ as long as $\tau_{j-1} \neq +\infty$. Then, there exists some finite J such that $\tau_J^b = +\infty$.

[13] This makes the analysis tractable, as other potentially payoff-relevant variables, such as the number of accumulated buyers and the (higher-order) beliefs of each of players about them, depend on players' *private histories*. Adding them would prevent the equilibrium from having a recursive structure.

accumulated buyer and that the seller's continuation play coincides with the one specified by the equilibrium strategy.[14] In the remainder of the paper, we call these, simply, equilibria.

2.1 Discussion of the Assumptions

To study price dynamics in a revenue-management environment without commitment, we make some rather unconventional modeling choices. These assumptions aim at giving tractability to a problem that otherwise has been proven intractable, while keeping the main trade-offs of economic interest. In particular, these assumptions allow us to prove that there exists a unique equilibrium, to characterize it using dynamic programming techniques, and to show that it exhibits price dynamics consistent with observed regularities. Before proceeding further, we discuss the role of these assumptions.

Inattention frictions. We assume that being attentive is costly for accumulated buyers. As we explain in the Introduction, this is a simple modeling device to lower the buyers' willingness to monitor the price, and therefore it limits the ability of the seller to influence the beliefs through deviations. It is necessary for making the model tractable, and it captures the observation that buyers check the price only occasionally. We also assume that buyers observe the regular offer at their arrival time for free. This assumption ensures that some goods are sold through regular offers in equilibrium, while the seller has a limited ability to manipulate the buyers' beliefs. In Section 4.2, we discuss some alternative ways to model buyers' attentiveness and argue that our main results are preserved.

Flash sales. Low-price flash sales are accessible to both accumulated buyers and shoppers in a broader market, while regular sales are accessible only to attentive buyers. This preserves the standard dynamic pricing trade-off of the seller, namely, sell fast at a low price, or slowly at a high price. Also, as an accumulated buyer can wait for and obtain a good in a flash sale, the seller is unable to fully extract surplus from buyers, making the dynamic price discrimination problem non-trivial.

We make a number of assumptions to facilitate the analysis. First, we assume that at each time t, the seller's decision on the number of goods offered in a flash sale depends on the number of regular sales at time t. Therefore, the seller is able to immediately respond to a buyer's purchase

[14] A rationale for this refinement of the buyer's off-the-path belief is that she interprets each observed deviating regular price as a "tremble" and the price is on the equilibrium path for the rest of the time. Since a buyer can pay attention for at most a zero-measure set of times, these trembles do not induce more accumulated buyers with positive probability. Our refinement is in the spirit of the so-called "passive beliefs" introduced by Hart and Tirole (1990). See McAfee and Schwartz (1994) and Horn and Wolinsky (1988) for its use in applications.

without a lag. This assumption implies that, in any strategy profile, when a buyer is deciding whether to accept a regular price, the time of the "next" flash sale is always well defined. The times of future flash sales are required to define the buyer's continuation payoff by rejecting the regular price in our continuous time model.

Second, we assume that, at each flash-sale time, an accumulated buyer observes the flash-sale offer only if she obtains a good in it. It ensures that, off the path, no buyer obtains information that allows her to make inferences about the number of accumulated buyers. This assumption, together with the one that accumulation is unobservable, permits the existence of an equilibrium where the payoff of the seller is independent of the previously accumulated buyers and the seller conditions only on the current time and inventory to make pricing and flash-sale decision.

Third, we assume that when a flash-sale offer is posted, goods will be immediately allocated.[15] This assumption simplifies the analysis, but our insight does not depend on this assumption. In the working paper version (Dilme and Li 2018), we consider an alternative model where it takes time to attract attention to flash-sale offers, and we show that our results are robust as flash-sale offers draw attention sufficiently quickly. This is also related to the assumption that the broader market is large and stationary enough to ensure that a seller can never exhaust the demand of flash sales from shoppers. Incorporating a history-dependent population of low-value shoppers would not only complicate the analysis without adding new insights, but also prevent the existence of Markov equilibria.

We make two additional simplifying assumptions, which can be relaxed without undermining our key results. We assume that consumers automatically accept a flash-sale offer. This assumption is made for simplicity. As it is suboptimal for the seller to charge a regular price lower than v_L, it would naturally be part of equilibrium behavior to accept such a low price even if buyers were allowed to reject the offer. In addition, we assume that holding a flash sale is costless. Assuming that offering such a deal involves a small cost $c' \in (0, v_L)$ would not qualitatively change our results although it would give the seller less incentive to lower the price.[16]

[15] In practice, the deal alert can be email/text messages sent by third parties, such as Kayak or Orbitz in the airline industry, or online real-time bidding (RTB) advertisements that facilitate the seller's ability to target and track "window shoppers." The time windows for these flash sale offers are typically very short compared to the length of the transaction season. See Biyalogorsky et al. (2001) for a discussion of a similar result in a different setting. We thank a referee for asking us to clarify this point.

[16] In a durable good setting, Öry (2017) argues that costly sales endow the seller with some commitment power and avoid the well-known Coase conjecture as the equilibrium outcome.

3 Analysis

We begin with a lemma that will simplify our analysis.

Lemma 1. *In any equilibrium,*

1. *the seller posts, in any history, a regular price that makes an attentive buyer indifferent between accepting and rejecting it, and*
2. *if a buyer does not purchase a good at her arrival time, she never rechecks the regular price.*

In equilibrium, an attentive buyer believes that there are no accumulated buyers. Also, she expects that the seller and the buyers who arrive in the future will follow a Markov strategy. Therefore, the continuation value she expects from rejecting the current regular offer depends only on the current state (t, k). Following the standard Diamond-paradox argument, the seller always finds it optimal to post the regular price that an attentive (either accumulated or newly arrived) buyer is indifferent between accepting and rejecting. This further implies that each buyer can achieve her equilibrium payoff by rejecting all equilibrium regular offers and waiting to possibly obtain a good in a flash sale. Then, rechecking is strictly suboptimal: it is costly and, after rechecking, it is optimal for the buyer to reject the equilibrium offer.

Lemma 1 implies that, in any equilibrium, if a buyer is accumulated off the path, she does not recheck the regular price. This generates three important observations. First, in equilibrium, at each time, a regular price offer is observed and accepted only by arriving buyers. Since the probability of two buyers arriving at the same instant is zero, the probability that two or more buyers observe a given price offer at the same time is zero too. Thus, the competition between buyers is inter-temporal rather than contemporaneous. Second, since accumulated buyers do not recheck the price, a buyer who rejects the regular price at her arrival time will not take any future regular offer; and thus she can obtain a good only in a future flash sale. Consequently, when a buyer arrives, she trades off between purchasing the good at the current regular price and waiting for a future flash sale. The possibility of obtaining a good in a future flash sale makes the continuation value she expects from rejecting an offer non-trivial. Finally, the lemma also implies that if a seller accumulates buyers off the path of play, his continuation payoff is independent of the number of accumulated buyers. This is because such an accumulation is not observed by other buyers, and accumulated buyers buy only in flash sales. As a result, by accumulating buyers, the seller can neither extract surplus from them nor increase the future buyers' willingness to pay. Consequently, in any equilibrium, the continuation payoff of the seller depends on only the time and the stock independently of the previous history.

3.1 Single-Unit Case

We first analyze the case where the seller has only one good, which helps to develop some intuition for the multi-unit model.

Proposition 1. *Suppose that $K = 1$. There is a unique equilibrium. In this equilibrium, at any time $t \in [0,1]$,*

1. *the regular price $p_1(t)$ satisfies*

$$v_H - p_1(t) = e^{-\lambda(1-t)}\beta(v_H - v_L)\,; \quad (1)$$

2. *an attentive buyer accepts a regular price offer if and only if it is weakly below $p_1(t)$, and never rechecks if she rejects an offer; and*

3. *the number of units offered in the flash sale is $f_1(t) = 0, \forall\, t < 1$, and $f_1(1) = 1$.*

Equilibrium timing of flash sales. Proposition 1 establishes that the seller holds a flash sale at time t (i.e., $f_1(t) = 1$) if and only if $t = 1$. The "if" part of the claim is a consequence of the seller's inability to commit: if the unit has not yet been sold at the deadline, the seller prefers to obtain some revenue $v_L > 0$ by offering a last-minute deal. This lowers the ability of the seller to extract rents from the buyers, as they know that they can wait until the deadline and obtain, with positive probability, the good at a low price. Still, due to the inter-temporal competition between buyers and the contemporaneous competition among buyers and shoppers at the deadline, the seller can keep the regular price above v_L. As a result, we obtain the "only if" part: the highest price accepted by a buyer before the deadline is strictly higher than v_L, and so it is optimal not to hold a flash sale before the deadline. Because a flash sale occurs only at the deadline, the good is allocated to a shopper only if no buyer arrives, making the allocation efficient. This will not be true in the multi-unit case.

Equilibrium regular pricing. The equilibrium price smoothly declines over time. The intuition is simple. The left-hand side of equation (1) represents a buyer's payoff if she arrives at time t and purchases the good at the equilibrium price $p_1(t)$, while the right-hand side represents the continuation value she expects from rejecting the offer. By Lemma 1, the equilibrium price $p_1(t)$ makes her indifferent between making the purchase upon arrival and rejecting the offer and waiting for the final flash sale. By waiting, she gets the good at the final flash sale only if (i) no other buyer arrives in $[t,1]$ and (ii) the good is assigned to her in the flash sale. The former event occurs with probability $e^{-\lambda(1-t)}$, which is increasing in t, and the probability of the latter event is β, conditional on the former event. Intuitively, as time passes, the probability that more buyers will arrive before the deadline shrinks, and so a buyer faces less competition and her reservation price decreases. The gap between $p_1(1) = (1-\beta)v_H + \beta v_L$ and v_L reflects the buyer's willingness to avoid competition with shoppers.

3.2 The Two-Unit Case

We now turn to the analysis of the multi-unit model. In this section, we focus on the case where the seller has two units. This is the simplest case where we can discuss the trade-offs that determine the timing of pre-deadline flash sales. In Section 3.3, we study the general model where the seller has more than two units.

Proposition 2. *Suppose that $K = 2$. There exists a unique equilibrium. There exists a $t_2^* \in [0,1)$ such that, in this equilibrium, at any $t \in [0,1]$ and $k \in \{1,2\}$,*

1. *the seller's regular price $p_k(t)$ satisfies (1) when $k=1$ and, when $k=2$,*

$$v_H - p_2(t) = \begin{cases} \int_t^{t_2^*} (v_H - p_1(s)) e^{-\lambda(s-t)} \lambda \mathrm{d}s \\ \quad + e^{-\lambda(t_2^*-t)} \big(\beta(v_H - v_L) + (1-\beta)(v_H - p_1(t_2^*))\big) & \text{if } t \le t_2^*, \\ \beta(v_H - v_L) + (1-\beta)(v_H - p_1(t)) & \text{if } t > t_2^*; \end{cases} \tag{2}$$

2. *an attentive buyer accepts a regular price offer at state (t,k) if and only if it is weakly below $p_k(t)$, and never rechecks if she rejects an offer; and*
3. *the number of units offered in a flash sale is $f_k(t) = 0$ if $t < t_2^*$, $f_k(t) = k-1$ if $t \in [t_2^*, 1)$, and $f_k(t) = k$ if $t = 1$.*

Proposition 2 shows that the equilibrium is characterized by a cutoff time t_2^*. In equilibrium, the seller holds no flash sale when $t \in [0, t_2^*)$. When $t \in [t_2^*, 1)$, instead, the seller holds a flash sale to ensure that his inventory size is not greater than one. That is, if $K_t = 2$ and no regular sale takes place, he sells one unit at the flash-sale stage; otherwise he holds no flash sale. At the deadline, the seller holds a flash sale for all remaining units. If the seller has k units at time t, the seller posts the highest regular price $p_k(t)$ such that if a buyer arrives she accepts the offer. In the remainder of this section, we shed light on why Proposition 2 holds.

Continuation play when $K_t = 1$. Given that buyers believe there is no accumulated buyer, the continuation payoff a buyer expects from rejecting an offer depends on the history only through the current time and inventory size. Also, the number of accumulated buyers does not affect the seller's continuation payoff: they obtain a good only in flash sales. As a result, both on and off the equilibrium path, once one of the units is sold, the continuation play is characterized by Proposition 1.

Equilibrium regular pricing when $K_t = 2$. The right-hand side of equation (2) represents a buyer's continuation value if she decides not to accept the current offer. There are two cases to consider. When $t \le t_2^*$, it has two terms. The first term corresponds to the event that one unit is sold at some time $t' \in [t, t_2^*]$ (because another buyer arrives). In this case, the inventory becomes 1 after the regular sale and therefore, by equation (1), her expected continuation payoff equals $v_H - p_1(t')$. The second term corresponds to the event that no buyer arrives in $[t, t_2^*]$. In this case, the buyer obtains the good at the time-t_2^* flash sale with probability β and, with probability $1-\beta$, she does not obtain a good and waits for the next flash sale. In the latter case, her continuation payoff equals $v_H - p_1(t_2^*)$.

When $t > t_2^*$ and $K_t = 2$, buyers expect an immediate flash sale. Thus, the continuation value of a buyer is expressed in equation (2) as the sum of her payoff if she gets the good at the time-t flash sale, $v_H - v_L$ and her continuation payoff if she does not get it, $v_H - p_1(t)$. The regular price $p_2(t)$ is then equal to the highest price that a buyer is willing to accept at time t when the inventory $K_t = 2$.

Equilibrium timing of flash sales. We now discuss how to obtain the equilibrium value of the flash sale time for the first unit, t_2^*. First, we rule out that the seller holds a two-unit flash sale before the deadline. To do so, fix some $t < 1$ and assume that $K_t = 2$. The payoff of the seller from holding a two-unit flash sale at time t is $2v_L$. The seller has the option of holding a one-unit flash sale at time t instead and obtaining a payoff equal to $v_L + \Pi_1(t)$, where

$$\Pi_1(t) \equiv \int_t^1 \lambda e^{-\lambda(t'-t)} p_1(t')\mathrm{d}t' + e^{-(1-t)\lambda} v_L \tag{3}$$

is, by Proposition 1, the expected profit for the seller at time t if $K_t = 1$. Since $\Pi_1(t) > v_L$ for $t < 1$ (notice that $p_1(t') > v_L$ for all $t' \in [0,1]$), holding a two-unit flash sale before the deadline is always dominated by holding a one-unit flash sale. Hence, in equilibrium, if there is a flash sale before the deadline, only one unit is offered in the sale.

We now determine the first time t_2^* where, if no regular sale has occurred before, the seller holds a flash sale. Assume that $t_2^* > 0$. We can compute the value of the seller from deviating and holding a flash sale at $t_2^* - \varepsilon$, for some small $\varepsilon > 0$. If no buyer arrives in $[t_2^* - \varepsilon, t_2^*]$, which happens with probability $e^{-\lambda\varepsilon}$, the payoff of the seller is not affected by the deviation. If, instead, exactly one buyer arrives in $[t_2^* - \varepsilon, t_2^*]$, which happens with probability $\varepsilon\lambda e^{-\lambda\varepsilon}$, the seller's gain from deviating is

$$v_L + p_1(t_2^*) - (p_2(t_2^*) + \Pi_1(t_2^*)) + O(\varepsilon) \tag{4}$$

as $\varepsilon \to 0$. For small ε, the event where more than one buyer arrives in $[t_2^* - \varepsilon, t_2^*]$ happens with probability $O(\varepsilon^2)$. Hence, the seller has no incentive to deviate from holding flash sales at t_2^* only if expression (4) is weakly negative.

Since $p_1(1) > p_2(1)$ and $\Pi_1(1) = v_L$, expression (4) is strictly positive when it is evaluated at $t_2^* = 1$. Therefore, in any equilibrium, $t_2^* < 1$ and the seller offers one unit during the flash sale at t_2^* if $K_{t_2^*} = 2$. We can then compute the seller's gain from deviating by holding a flash sale at $t_2^* + \varepsilon$, for some $\varepsilon > 0$, while charging the equilibrium regular price in $[t_2^*, t_2^* + \varepsilon]$. Again, the seller's gain from such a deviation is mostly determined by the event where one buyer arrives in $[t_2^*, t_2^* + \varepsilon]$. Also, by equation (2), $p_2(\cdot)$ is continuous at t_2^*, and the gain from such a deviation in such an event is exactly the negative of expression (4). Therefore, we conclude that in an equilibrium where $t_2^* \in (0,1)$, it must be that $v_L + p_1(t_2^*) - (p_2(t_2^*) + \Pi_1(t_2^*))$ is equal to 0, whereas it must be weakly positive if $t_2^* = 0$. Using equation (2), we can replace $p_2(t_2^*)$ in the previous equation and obtain that

$$\beta v_L + (1-\beta) p_1(t_2^*) - \Pi_1(t_2^*) = 0, \tag{5}$$

whenever $t_2^* \in (0,1)$. In the proof of Proposition 2 we show that equation (5) has at most one solution for $t_2^* \in (0,1)$. If a solution exists, it corresponds to the value of t_2^* in Proposition 2. If, instead, no solution exists, then $t_2^* = 0$ in equilibrium; that is, the seller holds a fire sale for one unit at $t = 0$.

Note that since $t_2^* < 1$, the equilibrium allocation rule is *inefficient*: a unit is allocated to low-value shoppers when it is still possible that multiple high-value buyers will arrive before the deadline.

3.3 The Multi-Unit Case

We now consider the general case, where K is an arbitrary natural number.

Proposition 3. *There exists a unique equilibrium. The equilibrium is characterized by a decreasing sequence of threshold times* $(t_k^*)_{k=1}^{K+1}$ *satisfying that* $t_1^* = 1$, $t_{K+1}^* = 0$, *and* $t_{k+1}^* < t_k^*$ *whenever* $t_k^* > 0$. *In this equilibrium, for each* $t \in [0,1]$ *and* $k \in \{1,\dots,K\}$,

1. *the regular price* $p_k(t)$ *satisfies*

$$v_H - p_k(t) = \begin{cases} \int_t^{t_k^*} (v_H - p_{k-1}(s)) e^{-\lambda(s-t)} \lambda \mathrm{d}s \\ \quad + e^{-\lambda(t_k^* - t)} \big(\beta(v_H - v_L) + (1-\beta)(v_H - p_{k-1}(t_k^*))\big) & \text{if } t \le t_k^*, \\ \beta(v_H - v_L) + (1-\beta)(v_H - p_{k-1}(t)) & \text{if } t > t_k^*, \end{cases} \tag{6}$$

   where we set $p_0(t) = v_H$ *for convenience;*

2. *an attentive buyer accepts a regular price offer at state* (t,k) *if and only if it is weakly below* $p_k(t)$, *and never rechecks if she rejects an offer; and*

3. *the number of units offered in the flash sale is* $f_k(t) = k - \bar{k}$ *if* $k > \bar{k}$ *and* $f_k(t) = 0$ *if* $k \le \bar{k}$, *where* $\bar{k}$ *is such that* $t \in [t_{\bar{k}+1}^*, t_{\bar{k}}^*)$.

Equilibrium timing of flash sales. The flash-sale strategy of the seller defines a time-dependent "ceiling" for his inventory. It specifies the upper bound on the inventory size, denoted by $\bar{K}_t$, that the seller is willing to keep at time t. It is determined by finding the value $\bar{k}$ such that $t \in [t_{\bar{k}+1}^*, t_{\bar{k}}^*)$. To see this, suppose that at time t, $r \in \{0,\dots,K_t - 1\}$ units are sold at the regular price, and so $K_t - r$ units remain at the flash-sale stage. The seller then checks whether $K_t - r$ is above $\bar{K}_t$. If so, he holds a flash sale for $K_t - r - \bar{K}_t$ units to ensure that only $\bar{K}_t$ units are left afterwards; otherwise, if $K_t - r \le \bar{K}_t$, the seller does not hold a flash sale. Figure 2 plots $\bar{K}_t$ for $t \in [0,1]$ in an example.

The ceiling of the inventory $\bar{K}_t$ is decreasing in time. This is true because, as the time gets close to the deadline, the expected number of future arrivals decreases. Furthermore, Proposition 3 establishes that $t_{k+1}^* < t_k^*$ whenever $t_k^* > 0$, implying that $\bar{K}_t$ decreases in steps of size 1. The logic behind this result is similar to the intuition provided above for the result that $t_2^* < t_1^* = 1$. Indeed, suppose that $t_{k+1}^* = t_k^* > 0$ for some $k > 1$; that is, the size of the jump down from the ceiling $\bar{K}_t$ at time $t = t_k^*$ is 2. In this case, the regular price at time $t_k^* - \varepsilon$, for $\varepsilon > 0$ small, would be

$$\begin{aligned} p_{k+1}(t_k^* - \varepsilon) &= \beta_2(v_H - v_L) + (1-\beta_2)(v_H - p_{k-1}(t_k^*)) + O(\varepsilon) \\ &< \beta(v_H - v_L) + (1-\beta)(v_H - p_{k-1}(t_k^*)) + O(\varepsilon) \\ &= p_k(t_k^* - \varepsilon), \end{aligned}$$

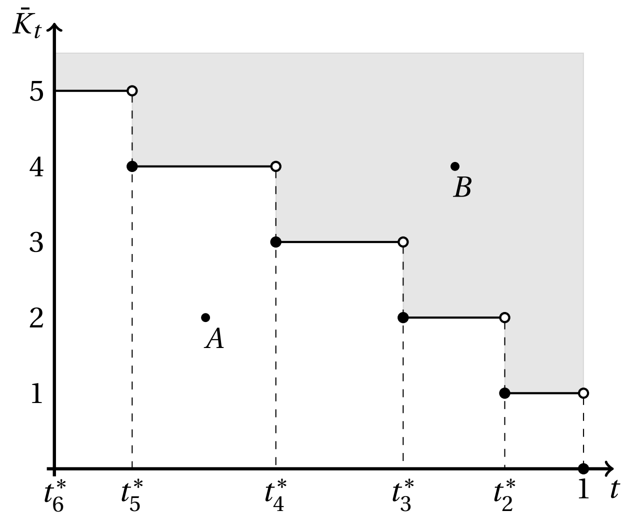

Figure 2: The step function $\bar{K}_t$ is the time-dependent ceiling for the seller's inventory. The seller does not hold a flash sale in states in the white area (including states on the boundary between the white and gray areas), such as at point A, whereas he holds an immediate flash sale in states in the shadowed area, such as at point B.

where, as before, $\beta_2 \equiv \beta + (1-\beta)\beta$ is the probability that an accumulated buyer obtains a good at a two-unit flash sale if there is no other accumulated buyer. Thus, similar to what we argued to show that $t_2^* < 1$, the seller is better off holding a flash sale for one unit at time $t_k^* - \varepsilon$ than following the equilibrium strategy. Intuitively, the seller can only *discretely* adjust his inventory by holding flash sales, while his incentive of doing so is determined by the distribution of arrivals in the remaining time, which changes *continuously* over time. Hence, the seller finds it optimal to lower the ceiling for his inventory unit by unit.

The equilibrium flash-sale strategy implies that, on the path of play, a multi-unit flash sale takes place only at $t = 0$ in the case where $K > \bar{K}_0 + 1$. After that, the seller still holds flash sales over time when realized demand is low, but only one unit is offered in each flash sale.

Equilibrium regular pricing. Proposition 3 establishes that, in a state (t, k) with $t \leq t_k^*$ and $k \geq 2$, the seller offers a regular price $p_k(t)$ that makes buyers indifferent whether to accept it. He does so until either there is a regular sale or t_k^* is reached and he holds a flash sale for one unit. As in the two-unit case, $p_k(t)$ coincides with the buyer's reservation price, given by equation (6).

If, instead, $t_{k'+1}^* \leq t < t_{k'}^*$ for some $k' < k$, then the seller immediately holds a flash sale for $k - k'$

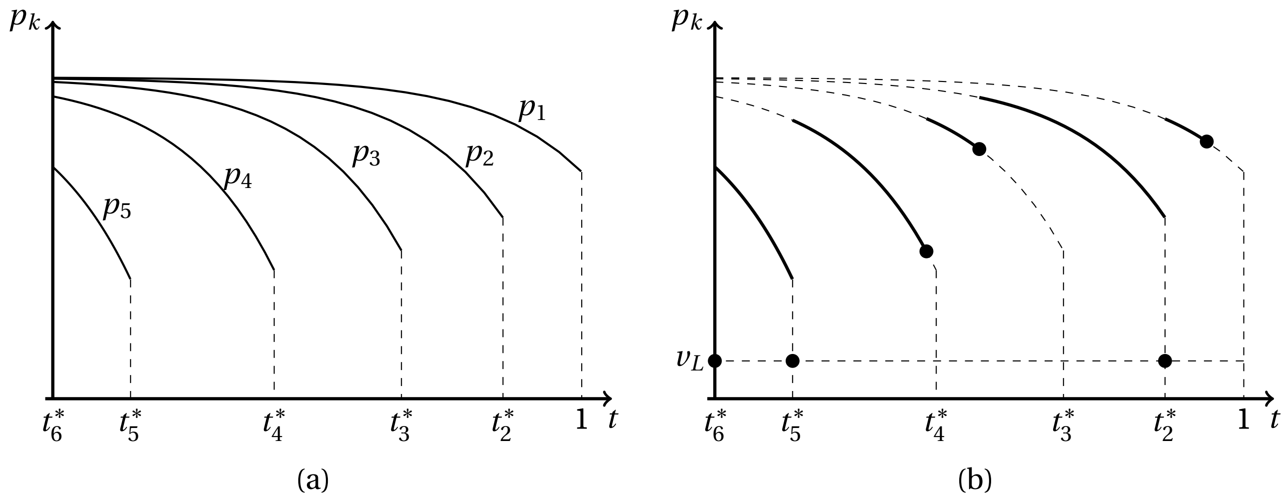

Figure 3: (a) depicts $p_K(\cdot)$ in $[0, t_K^*]$ for different values of K. (b) depicts a possible realization of the price path, where the dots denote transaction prices.

units. As a result, in this case, the following equation holds:

$$v_H - p_k(t) = \beta_{k-k'}(v_H - v_L) + (1 - \beta_{k-k'})(v_H - p_{k'}(t)) \,, \tag{7}$$

where $\beta_{k-k'}$ denotes the probability that a given buyer in the waiting state obtains the good at time t if $k-k'$ goods are offered and she is the only buyer in the waiting state at time t. Notice that $\beta_{k-k'}$ is increasing in both β and $k-k'$, for all $k-k' \geq 1$. When $k-k' = 1$, equation (7) takes exactly the same form as the one in (6). When $k-k' > 1$, equation (7) also applies when, while keeping k' the same, k is replaced by $k-1$, and so $v_H - p_k(t) = \beta(v_H - v_L) + (1-\beta)(v_H - p_{k-1}(t))$ and (6) holds.

Figure 3 provides some simulated price paths. Conditional on the inventory size, the price declines as the deadline approaches. However, transactions occur over time, and each transaction triggers an uptick in the price. As a result, the price goes up or down depending on the arrival of buyers.

3.4 The Role of the Competition between Consumers

Our model has two kinds of consumers: the (high-value) buyers and the (low-value) shoppers. The equilibrium price dynamics is affected by the degree of competition between consumers, which is determined by two parameters in our model: the arrival rate of buyers, λ, and the intensity of the competition in a flash sale, β. We devote this section to investigating how the change in the competition between consumers affects the timing of flash sales. These comparative statics exercises offer empirical testable implications.[17]

The variation of λ. The following result illustrates the effect that changes in λ have on the timing of flash sales.

Proposition 4. *The flash-sale time t_k^* increases in λ for every $k = 1, \ldots, K$.*

To understand the logic behind Proposition 4, notice that in our model the arrival rate of buyers sets the pace of the game: at a given instant t, the future arrivals from t to 1 are distributed according to a Poisson distribution with parameter $(1-t)\lambda$. Thus, for a larger λ, the seller expects that more buyers will arrive in the future after any given time, and so he has a lower incentive to hold a flash sale, and flash sales are closer to the deadline.[18]

Given Proposition 4, the following implication is immediate.

Corollary 1. *The expected number of flash sales is decreasing in λ.*

The result is straightforward since, when λ is high, not only does the seller delay flash sales but also more buyers arrive in expectation.

The variation of β. To see the role of β on the equilibrium outcome, we analyze the case where β is either large or small. First, we consider the case where β is close to 1. In this case where, off the path, there is one accumulated buyer and the seller holds a flash sale, the buyer obtains a good with high probability. The following result shows that the seller sells most of the goods through an initial flash sale due to his inability to commit to future high prices:

Proposition 5. *There exists a unique $\bar{\beta} \in (0,1)$ such that $t_k^* = 0$ for all $k > 1$ if and only if $\beta \in [\bar{\beta}, 1)$. Furthermore, $\bar{\beta}$ is increasing in λ.*

Proposition 5 establishes that, when $\beta \geq \bar{\beta}$, the seller holds a flash sale at time 0 for $K-1$ units and keeps only one unit after that. That is, when buyers are very likely to obtain a good in a flash sale, the seller can generate only inter-temporal competition for one unit in equilibrium. To obtain further intuition, assume by way of contradiction that $t_2^* > 0$ in the extreme case where $\beta = 1$. Consider a history such that $K_{t_2^*-\varepsilon} = 2$ for some small and positive ε. In this case, the large

[17] Aryal et al. (2018) document the passenger heterogeneity and stochastic demand within and across flights in the airline industry.

[18] A similar argument implies that when the arrival rate is a time-dependent function $\lambda(\cdot)$, our equilibrium outcome remains qualitatively the same. The time variable can be "stretched" so that, under the normalized time, the arrival rate is constant.

value of β lowers the willingness to pay of a buyer who is attentive at $t_2^* - \varepsilon$: she knows that if she rejects the offer, she is very likely to obtain the good in a flash sale as the probability that another buyer arrives in $(t_2^* - \varepsilon, t_2^*]$ is $O(\varepsilon)$. Therefore, $p_2(t_2^* - \varepsilon) = v_L + O(\varepsilon)$, which implies that, conditional on one buyer arriving in $(t_2^* - \varepsilon, t_2^*]$, the payoff of the seller is given by

$$v_L + e^{-\lambda(1-t_2^*)}\mathbb{E}\big[\, p_1(t) \,\big|\, t > t_2^* \,\big] + \big(1 - e^{-\lambda(1-t_2^*)}\big) v_L + O(\varepsilon) \tag{8}$$

as $\varepsilon \to 0$, where the expectation is over time t when the next buyer arrives. The seller can instead hold a flash sale at $t_2^* - \varepsilon$. In this case, if a buyer arrives in $(t_2^* - \varepsilon, t_2^*]$, the payoff of the seller is $v_L + p_1(t_2^*) + O(\varepsilon)$, which is higher than the payoff in (8) for ε small enough, given that $p_1(\cdot)$ is decreasing. Therefore, conditional on one buyer arriving in $(t_2^* - \varepsilon, t_2^*]$, the seller prefers holding a flash sale at $t_2^* - \varepsilon$ rather than at t_2^*. If, instead, no buyer arrives in $(t_2^* - \varepsilon, t_2^*]$, the deviation does not change the seller's payoff. Finally, the event where two or more buyers arrive in $(t_2^* - \varepsilon, t_2^*]$ is highly unlikely (its probability is $O(\varepsilon^2)$). Then, the expected gain from deviating is positive, which indicates that there is no equilibrium with $t_2^* > 0$.

The value of $\bar{\beta}$ is found by requiring that (5) hold when $t_2^* = 0$ and $\beta = \bar{\beta}$. One can show that t_2^* is decreasing in β, and so, for $\beta > \bar{\beta}$, $t_k^* = 0$ for all $k \geq 2$. Intuitively, a higher β increases a buyer's payoff from waiting for the next flash sale, resulting in a lower equilibrium regular price for each state (t, k). Furthermore, it increases the price difference between holding one and two units at each time t, $p_1(t) - p_2(t)$, and therefore strengthens the seller's incentive to hold a flash sale, as we discussed in Section 3.2. Notice that $p_1(t) - p_2(t)$ is decreasing in t, and so a larger value of β implies that, in equilibrium, the seller holds the first flash sale earlier; that is, t_2^* decreases in β. Using Proposition 4 and that t_2^* is decreasing in β, it is easy to see that $\bar{\beta}$ is increasing in λ.

In the limit as $\beta \to 0$, the probability that a buyer obtains a good in a flash sale vanishes, and so the buyer's willingness to pay increases to v_H. As a result, buyers behave like myopic players as in Gallego and Van Ryzin (1994), and the seller's profit achieves its first-best value.

Proposition 6. *As $\beta \to 0$, $t_k^* \to 1$ and $p_k(t) \to v_H$ for any $k = 1, \ldots, K$ and $t \in [0, 1]$.*

The parameter β captures the likelihood that accumulated buyers notice a flash sale. In practice, it is determined by multiple factors. For example, it is likely to depend on the effective demand from people who consider purchasing the good only if it is cheap enough. In the airline example, the demand from the broader market may be constituted, in part, by the leisure passengers who search for low-price tickets for vacations, and so one can expect β to be lower during the vacation season. The value of β is also likely to depend on the ability (technology) of sellers or third-party agencies to attract accumulated buyers' attention. Nowadays, it has become easier for airlines and price aggregators—like Expedia, Kayak or Orbitz—to track accumulated buyers (window shoppers) by using "cookies," that enable them to send deal-alert emails or text messages, or bid on personalized online advertisements in real time. Although the advance of information technologies gives the airlines and third parties more flexibility to distinguish between and price discriminate against consumers, it may have an ambiguous effect on β. On the one hand, the availability of big data enables third-party agencies to more precisely tailor and lure back high-value buyers and prioritize their purchases to build long-run customer capital, which increases β.[19] On the other hand, it makes it easier for the seller to target shoppers separately from buyers by, for example, holding flash sales at inconvenient times for buyers. In addition, the development of information technologies also makes it easier for low-value shoppers to react to flash sales, which may decrease β. The fact that, by Proposition 6, the seller benefits from low values of β rationalizes the observation that in the airline industry deal-alert services are provided often by third parties instead of airlines, even though the airline could do so at a lower cost.

4 Further Discussion

This section discusses some possible extensions of our model.

4.1 Multiple Types

We assume that buyers have the same valuation for the good, v_H, so that they only differ in their private histories. This assumption ensures that buyers purchase immediately upon arrival in equilibrium, which is necessary to keep (t, K_t) as the only state variable. The existence of equilibria without the accumulation of buyers result is robust to perturbing the heterogeneity in buyers' value: as long as the valuations of the buyers are not too dispersed, there is a Markov equilibrium where the regular price is immediately accepted by all types of buyers. The following result states that, when the distribution of valuations is uniform, as long as the buyers' valuation is bounded away from v_L and is not too sparse, there are equilibria as described in Section 3.

[19] Öry (2017) assumes that it is costly to attract accumulated buyers' attention and argues that such a cost has been lowered with the development of Internet. In this sense, the cost of attracting accumulated buyers in Öry (2017) has a similar flavor to the $1-\beta$ in our model: they both capture the inefficiency that endows the seller some commitment power. Relatedly, Dana and Orlov (2014) find that airline capacity utilization has dramatically increased since 1993, and they argue that the popularization of Internet has much to do with it.

Proposition 7. *Fix $\underline{v_H} > v_L$. There exists some $\overline{v_H}^* > \underline{v_H}$ such that if the types of buyers are uniformly distributed in $[\underline{v_H}, \overline{v_H}]$ for $\overline{v_H} \in (\underline{v_H}, \overline{v_H}^*)$ then there is an equilibrium without accumulation as described in Proposition 3.*

The logic behind Proposition 7 is the following. In a static model where a monopolist makes a take-it-or-leave-it offer to a buyer, if the buyer's valuation is uniformly distributed in $[\underline{v_H}, \overline{v_H}]$ and $\overline{v_H} \leq 2\underline{v_H}$, it is optimal for the seller to charge a price equal to $\underline{v_H}$. Indeed, increasing the price above $\underline{v_H}$ generates a decrease in the probability of trade that is large enough that the monopolist prefers to ensure trade by setting the price equal to $\underline{v_H}$. The intuition is similar in our dynamic model: increasing the price above the reservation price of the buyers with the lowest valuation effectively implies losing them as potential buyers, since they will not recheck the price in the future. Even though increasing the price increases the revenue if a buyer with a high valuation arrives, the negative effect dominates the positive one, as in the static model. This result can be generalized to distributions of types with a support bounded away from v_L and with a probability density function bounded away from 0.

A higher buyer heterogeneity would prevent the existence of Markov equilibria, making the model intractable. In this case, all equilibria would feature a history-dependent (and payoff-relevant), stochastic stock of accumulated buyers, and so the buyers' strategies would depend on their private histories. However, we conjecture that the key incentive of our model remains in a more general setup. Since, in general, one can expect the price to be low when the stock of goods is high, the seller has the incentive to further lower the price (or hold a flash sale) after a low realized demand to generate scarcity and increase the future buyers' willingness to pay. The lack of commitment of the seller is then likely to accelerate the price decay at some isolated times, and generate price dynamics similar to the ones in our model.

4.2 Attention Cost

We assume that buyers observe the regular offer at their arrival time for free, while it is costly to recheck it afterwards. This assumption simplifies our analysis by ensuring that some goods are sold through regular offers in equilibrium, while the seller has a limited ability to manipulate the buyers' beliefs. Other assumptions give similar results. For example, one can assume that the attention cost is zero, while keeping the assumption that buyers can choose to sample the regular price offer for a finite but arbitrarily large number of times. It is easy to verify that our equilibrium remains an equilibrium under this assumption. The reason is that, in our equilibrium, the buyers' gain from checking the price and stock is zero: independently of what she observes, she is indifferent between accepting the offer and rejecting it. Thus, fixing the seller's strategy, not rechecking the price is optimal for a buyer even if paying attention is costless. Once Lemma 1 is proven to hold, it is clear that the rest of the equilibrium analysis of the new model is identical to the equilibrium analysis in our base model.

Alternatively, one could assume that, once a buyer arrives, she is not automatically attentive at her arrival time. She can choose her first (free) attention time, but her subsequent attention times are chosen at a cost. Also, assume that during a flash sale, a good is assigned to an accumulated buyer with positive probability even if she has not used her first free attention time yet. In this model, one can show that there is an equilibrium where buyers choose their free opportunity to pay attention at their arrival time, and the equilibrium regular price and the timing of the flash sales are identical to the ones in our model. Intuitively, in our equilibrium, each buyer takes both the future arrival of other buyers and the dynamics of the regular price into account, making her indifferent to sampling the price at any moment, and thus she has no incentive to delay the first (free) sample upon her arrival. Notice that the previous observation does not imply that in our baseline model a buyer' expected payoff is independent of her any arrival time. Indeed, a buyer who arrives later in the game cannot get the good in some early flash sales in which she could otherwise obtain positive surplus.[20]

4.3 Observability of Inventory

We assume that the remaining inventory K_t is observable. This assumption is necessary to avoid a signaling problem: if the seller has payoff-relevant private information, he can use prices to signal this information, and this results in the standard multiplicity of equilibria and off-the-equilibrium-path beliefs. It is worthwhile to point out, however, that a simple unraveling argument shows that if it is costless and verifiable, the seller always reveals the remaining stock.

In practice, buyers may observe imperfect but informative signals of the inventory in some markets, providing the seller with some incentive to hold a flash sale. For example, in the airline industry, airlines sometimes report online the number of remaining available seats. Even though the advertised number of available seats may not coincide with the actual inventory (airlines sometimes block some seats for elite passengers), it is an informative proxy. Escobari (2012) uses the advertised number of available seats as a proxy for the real inventory and empirically studies the price patterns in the airline industry. He finds that the price significantly increases as

[20]For example, it is easy to see that, in the case $K = 2$, the equilibrium expected payoff of a buyer arriving in $[0, t_2^*]$ is independent of the particular arrival time $t \in [0, t_2^*]$, and this is also true for a buyer arriving in $(t_2^*, 1]$. Nevertheless, a buyer obtains a strictly higher equilibrium expected payoff if she arrives in $[0, t_2^*]$ than if she arrives in $(t_2^*, 1]$.

the number of available seats decreases. This suggests that the number of available seats is a signal of inventory that is interpreted as credible. Williams (2018) analyzes a new airline dataset that distinguishes between blocked and occupied seats. He argues that seat maps observed by buyers are a useful proxy for bookings, even though they overstate the latter by approximately 10%.

4.4 Creating Scarcity by Discarding Inventory

In our model, the seller sometimes finds it optimal to create scarcity by holding flash sales. Holding a flash sale is profitable for the seller because it reduces the available stock and allows him to increase the price in the future. Nevertheless, because the seller cannot commit to the timing of the flash sales and buyers may obtain goods through them, the prospect of future flash sales lowers the reservation value of the buyers and also their willingness to pay. Thus, after some histories, the seller would like to commit to not selling some of the units (or, equivalently, discarding them) instead of holding flash sales, which would allow him to effectively lower the stock without lowering the willingness to pay of the current buyers.

Our results do not change if the seller has access to other mechanisms, in addition to flash sales, that allow him to lower the stock of goods as long as they generate a lower revenue than flash sales do. For example, suppose that the seller can reduce the stock by giving goods away in special promotions and obtaining zero revenue. Given that the seller lacks commitment power and has access to a market of shoppers, if there were an equilibrium where, after some histories, the seller would give some goods away, he would rather hold flash sales, as this would give him a higher revenue.

4.5 Returned Goods and Overbooking

In our model, transactions are irreversible: the inventory K_t weakly declines over time. In reality, however, some mechanisms make the stock of goods occasionally increase. One such possibility is when buyers regret buying the good after the purchase and ask to return it. Imagine that, in our model, returns occur with positive probability and, in such an event, K_t increases. In this case, as long as the probability of a good being returned is independent of the value that its buyer attached to it (at the purchase time), it is easy to accommodate our equilibrium construction to the possibility of returning goods. The reason is that, in this case, (t, K_t) remains the state variable of a Markov equilibrium and the equilibrium flash-sale timing is determined by the corresponding ceiling function $\bar{K}_t$: if a unit is returned at time t and the inventory becomes higher than $\bar{K}_t$, the seller immediately holds a flash sale to get rid of the extra unit; otherwise, the seller holds no flash sale at t and adjusts the regular price accordingly. Note that the possibility of these stock increases (and corresponding random flash sales) lowers the buyers' willingness to pay; as a result, holding flash sales is less effective to enhance the future regular price.

Similarly, the inventory K_t can go up if the seller was allowed to overbook and buy back sold goods. In this case, the seller may endogenously buy back sold goods from low-value buyers and sell it to high-value buyers. See Ely et al. (2017) for a discussion of this topic. Obviously, in this case, the flash sale is less effective in increasing the buyers' willingness to pay, but our mechanism remains as long as such reallocation is sufficiently costly.

4.6 On Commitment and Welfare

The outcome of the unique equilibrium described in Proposition 3 is inefficient when $K > 1$. Indeed, since the seller holds flash sales before the deadline with positive probability, some goods may be allocated to shoppers early in the game. This may result in rationing at later times, as it may be that more buyers arrive than the number of units left. Proposition 5 illustrates that such inefficiency is extreme when β is large: in this case, $K-1$ goods are purchased by (low-value) shoppers, rather than arriving (high-value) buyers.[21]

The (misallocation) inefficiency is mainly due to the lack of commitment by the seller. In our model, holding flash sales is the only credible way to reduce the buyers' continuation value and to raise their current willingness to pay. If the seller were allowed to publicly commit to a selling strategy, as in the so-called Stackelberg game, he would have more flexibility to obtain a higher surplus from trade (e.g., discarding goods or committing to holding flash sales only at the deadline) and, therefore, he would hold flash sales less frequently.[22] Characterizing the optimal Stackelberg pricing strategy of the seller is beyond the scope of this paper. Instead, we discuss an example of a strategy where the seller holds flash sales only at the deadline, leading to an efficient allocation. The example illustrates that, under certain conditions, the seller is better off committing to efficiently allocating goods to low-value shoppers only at the deadline.

Consider the case $K = 2$. Suppose also that the seller commits to holding flash sales only at the deadline (for all remaining units) and to offering a regular price that makes attentive buyers indifferent whether to buy a good. Suppose that buyers buy upon arrival. Lemma 1 still holds: if a buyer rejects the offer, she obtains a good only in a flash sale. When $K_t = 1$, the regular price is

[21] As Dana (1998) points out, misallocation across consumers is typical when prices are rigid. Our result indicates that it can take place even if the price is optimized by the seller at each moment.

[22] This is in line with the prediction of Board and Skrzypacz (2016). When the seller can commit to a more general dynamic mechanism, he has no incentive to hide his inventory.

given by equation (1). When $K_t = 2$, then the regular price $\bar{p}_2(t)$ satisfies

$$v_H - \bar{p}_2(t) = \underbrace{e^{-\lambda(1-t)}\beta_2(v_H - v_L)}_{\text{no buyer arrives}} + \underbrace{e^{-\lambda(1-t)}\lambda(1-t)\beta(v_H - v_L)}_{\text{one buyer arrives}}, \tag{9}$$

where the right-hand side of the previous equation represents the continuation value a buyer expects if she rejects the regular offer. If no other buyer arrives before the deadline, the seller holds a two-unit flash sale at the deadline, and the accumulated buyer gets a good with probability $\beta_2 \equiv \beta + (1-\beta)\beta$. If exactly one buyer arrives in $[t, 1]$, the seller holds a one-unit flash sale at the deadline, and the accumulated buyer gets a good with probability β. If multiple buyers arrive, the accumulated buyer does not obtain a good. Therefore, at time t, the buyer is willing to accept a regular price p if and only if $p \leq \bar{p}_2(t)$. One can show that by following this strategy, the seller's expected profit at $t = 0$ is strictly greater than his equilibrium payoff as long as λ is not too small. Indeed, in the equilibrium described in Proposition 2, before t_2^*, the price that the seller can charge decreases as t_2^* approaches. By postponing the flash sale to the deadline, for each t when $K_t = 2$, the seller increases the buyers' reservation price from $p_2(t)$ to $\bar{p}_2(t)$, without affecting the continuation equilibrium when $K_t = 1$. Nevertheless, it is not the case that postponing the flash sale generates a higher payoff to the seller in all events. Consider, for example, the event where exactly one buyer arrives, and she arrives at a time $t \in (t_2^*, 1]$. When flash sales occur only at the deadline, the seller's payoff is $\bar{p}_2(t) + v_L$, which is strictly lower than the equilibrium payoff, $v_L + p_1(t)$. Thus, a later flash sale increases the willingness to pay of the buyers when $K_t = 2$, but lowers the probability that a buyer purchases a good when $K_t = 1$. When λ is not small, the high likelihood that multiple buyers arrive early in the game favors holding a flash sale only at the deadline. Hence, when the seller has some commitment power, he may find it attractive to delay or even abandon flash sales. In doing so, he can charge a higher regular price when t is small and $K_t = 2$.

A Proofs

A.1 Proof of Results prior to Section 3.2

Proof of Lemma 1. The proof has three steps.

Step 1. We show that the continuation value that an attentive buyer expects from rejecting a regular price at some time t depends only on t and the remaining stock (and not on his private history or the particular on- or off-the-path regular price at the current time). To see this note first that in an equilibrium, a buyer believes that there is no accumulation regardless of her private history. A buyer who is attentive at time t assigns a probability of zero to the event that another (newly arrived or accumulated) buyer observes the price offer. Hence, the buyer believes that if she rejects the offer, the probability that there is no transaction at the regular price at time t is 1. Also, she expects the seller's continuation strategy to be Markov and hence independent of the current regular price. Therefore, at any state (t, k), the continuation value that an attentive buyer expects from rejecting the regular offer depends on her history only through (t, k), and it is denoted by $u_k(t)$.

Step 2. We show that in any equilibrium, the equilibrium price $p_k(t)$ at each state (t, k) makes buyers indifferent whether to accept it or not; that is, $v_H - p_k(t) = u_k(t)$. Indeed, offering a price strictly lower than $v_H - u_k(t)$ is clearly suboptimal, since the seller can slightly increase such a price while guaranteeing that it will be accepted. Conversely, if the equilibrium price is higher than $p_k(t)$, then buyers will reject the offer for sure; however, this is inconsistent with the equilibrium condition that buyers believe that there is no accumulation.

Step 3. We show that in any equilibrium, it is strictly suboptimal for an accumulated buyer to recheck the price. Assume to the contrary that there is an equilibrium where off the path, a buyer that is accumulated pays attention at some future time. Consider then a deviation consisting of rejecting all equilibrium offers while following the equilibrium strategy regarding the rechecking times. Under such a strategy, the buyer never buys a good through a regular offer, and she can obtain a good only in flash sales. Also, since she is indifferent between accepting and rejecting equilibrium offers, this strategy gives the buyer the same payoff as the equilibrium strategy. Now, consider a third strategy where an accumulated buyer never pays attention. This strategy dominates the second strategy since the buyer saves the attention cost while keeping the same probability of getting the good in a flash sale. Consequently, an accumulated buyer finds it strictly suboptimal to pay attention in any equilibrium.

□

Proof of Proposition 1. Note first that in any equilibrium $f_1(1) = 1$; that is, the seller holds a flash sale to sell the remaining unit at the deadline. This gives him a payoff of v_L instead of a payoff of 0.

Fix an equilibrium and a time t, let $\tau_1(t) \in [t, 1]$ be the next time when a flash sale happens according to the equilibrium strategy. Formally,

$$\tau_1(t) = \inf\{t'' \in [t, 1] | f_1(t'') = 1\}.$$

Note that $\tau_1(t') = \tau_1(t)$ for all $t' \in [t, \tau_1(t)]$. This is the case because if $\tau_1(t) > t$, then, by the definition of $\tau_1(t)$, $f_1(t'') = 0$ for all $t'' \in [t, \tau_1(t))$. Hence, for any $t' \in [t, \tau_1(t))$, $f_1(t'') = 0$ for all $t'' \in [t', \tau_1(t))$.

By Lemma 1, the equilibrium regular price at time t, denoted by $p_1(t; \tau_1(t))$, is such that the payoff a buyer obtains from purchasing is equal to the payoff she expects from rejecting the equilibrium price; that is,

$$v_H - p_1(t; \tau_1(t)) = e^{-\lambda(\tau_1(t) - t)} \beta (v_H - v_L) \, .$$

The first argument of $p_1(\cdot; \cdot)$ corresponds to the calendar time, and the second one corresponds to the time a buyer expects the next flash sale to happen if there is no regular sale before then. If a buyer rejects a regular price at time t, she obtains the good with probability β if no other buyer arrives in $(t, \tau(t)]$. Note that $p_1(\cdot; \tau_1(t))$ is increasing in $[t, \tau_1(t)]$. Therefore,

$$p_1(t; \tau_1(t)) \geq p_1(\tau_1(t); \tau_1(t)) = (1 - \beta) v_H + \beta v_L > v_L \, .$$

The corresponding payoff of the seller at time t is

$$\Pi_1(t; \tau_1(t)) = \overbrace{\int_t^{\tau_1(t)} e^{-\lambda(t' - t)} p_1(t'; \tau_1(t)) \lambda \mathrm{d} t}^{\text{buyers arrive in} [t, \tau_1(t)]} + \overbrace{e^{-\lambda(\tau_1(t) - t)} v_L}^{\text{no buyer arrives}} \, .$$

Now we prove that there is no equilibrium where $\tau_1(t) < 1$ for some $t < 1$. Suppose to the contrary that there exists such an equilibrium. This implies that the payoff of the seller at time $\tau_1(t)$ is $\Pi_1(\tau_1(t); \tau_1(t)) = v_L$. If the seller instead deviates by charging the price $p_1(t'; \tau_1(t')) > v_L$ for all $t' \geq t$ (which is accepted by arriving buyers) and holding a flash sale only at the deadline (if no buyer has arrived), he obtains

$$\int_t^1 e^{-\lambda(t'-t)} p_1(t';\tau_1(t'))\lambda \mathrm{d}t' + e^{-\lambda(1-t)} v_L > v_L \,.$$

This is a contradiction. Therefore, in any equilibrium, it must hold that $\tau_1(t) = 1$, for all $t \in [0,1]$, and that $p_1(t) = p_1(t;1)$. Therefore, the regular price satisfies equation (1).

We can now verify that no player has a profitable deviation in the candidate equilibrium where $\tau_1(t) = 1$, for all t. Buyers have no incentive to deviate since the equilibrium regular price makes them indifferent whether to reject the offer. Consider the seller's incentive to deviate. Fix a time t. Consider the deviation for the seller consisting of setting some $t_1 \in [t,1)$ for the next flash sale and charging some price $\tilde{p}_1(t')$ at each time $t' \in [t,1]$. Let B be the set of times $t' \in [t,t_1]$ where $\tilde{p}_1(t) \le p_1(t)$; then buyers accept the price at time t' if and only if $t' \in B$. Therefore, the seller's continuation payoff from the deviation is given by

$$\tilde{\Pi}_t \equiv \int_t^{t_1} e^{-\int_t^{t'} \lambda \mathbb{I}_B(t'')\mathrm{d}t''} \tilde{p}_1(t')\mathbb{I}_B(t')\lambda \mathrm{d}t + e^{-\lambda\mu(B)} v_L \,,$$

where $\mathbb{I}_B(\cdot)$ is the indicator function and μ is the Lebesgue measure in $\mathbb{R}$. Since $p_1(\cdot)$ is decreasing and strictly higher than v_L, it is clear that the seller is better off setting $\tilde{p}_1(t) = p_1(t)$ for all t and $t_1 = 1$. More formally,

$$\begin{aligned}
\tilde{\Pi}_t &\le \int_t^{t_1} e^{-\int_t^{t'} \lambda \mathbb{I}_B(t'')\mathrm{d}t''} p_1(t')\mathbb{I}_B(t')\lambda \mathrm{d}t + e^{-\lambda\mu(B)} v_L \\
&\le \int_t^{t+\mu(B)} e^{-\lambda(t'-t)} p_1(t')\lambda \mathrm{d}t + e^{-\lambda\mu(B)} v_L \\
&\le \int_t^{1} e^{-\lambda(t'-t)} p_1(t')\lambda \mathrm{d}t + e^{-\lambda(1-t)} v_L = \Pi_1(t;1) \,,
\end{aligned}$$

where $\Pi_1(t;1) \equiv \Pi_1(t)$ is the seller's equilibrium payoff at time t. Therefore, the seller does not gain from the deviation. □

A.2 Proof of Results in Section 3.2

Proof of Proposition 2. We first obtain some necessary conditions satisfied by any equilibrium. Then we verify that these conditions are satisfied by a unique equilibrium. Notice that, in any equilibrium, if at some time $t \in [0,1]$, $K_t = 1$, then the continuation play is given by Proposition 1. Indeed, given our focus on equilibria where a buyer believes that there is no accumulation, the reservation value of a buyer arriving at t and observing $K_t = 1$ satisfies equation (1) and, therefore, the seller finds it optimal to charge a regular price $p_1(t)$ and to set $f_1(t) = 0$ for all $t < 1$ and $f_1(1) = 1$.

Fix an equilibrium. For each time t we use $\tau_2(t) \geq t$ to denote the next time when a flash sale occurs if $K_t = 2$ and no regular sale takes place. That is,

$$\tau_2(t) \equiv \inf\{t' \geq t | f_2(t') > 0\} \in [t, 1] \cup \{\infty\},$$

where we use the convention $\inf \emptyset = \infty$. Notice that, as we argued for τ_1 in the proof of Proposition 1, if $\tau_2(t) > t$ and $t' \in [t, \tau_2(t)]$, then $\tau_2(t') = \inf\{t'' \geq t' | f_2(t'') > 0\} = \tau_2(t)$.

We divide the proof into two steps. The first step sets some properties that $\tau_2(\cdot)$ must satisfy in equilibrium, and the second step establishes that (i) there is a unique τ_2 which satisfies these properties, and (ii) there is a unique equilibrium.

Step 1. *In any equilibrium, the seller holds a flash sale for two units only at the deadline, i.e., $f_2(t) = 2$ if and only if $t = 1$.*

Proof. At time $t = 1$ it is clear that the seller holds a flash sale for all remaining units, i.e., $f_2(1) = 2$.

Assume by way of contradiction that $f_2(t) = 2$ for some $t < 1$. Notice that if $K_t = 2$, the payoff of the seller at t is $2v_L$. If the seller deviates and holds a flash sale for only one unit, he obtains a payoff of $v_L + \Pi_1(t) > 2v_L$ where Π_1 is given in equation (3), a contradiction. □

Before proceeding, we introduce some notation. In any equilibrium, the continuation value a buyer expects from rejecting the offer at state $(t, 2)$ with $t < 1$ can be written in terms of the next flash-sale time $\tau_2(t) \in [t, 1]$. Indeed, if another buyer arrives at $t' \in [t, \tau_2(t)]$, then by equation (1), the equilibrium continuation value of the buyer equals $v_H - p_1(t')$ at time t', whereas if no buyer arrives in $[t, \tau_2(t)]$, her expected equilibrium payoff at time $\tau_2(t)$ is

$$\beta(v_H - v_L) + (1 - \beta)(v_H - p_1(\tau_2(t))) \, .$$

Thus, for each $t \in [0, 1]$ and $t_2 \in [t, 1]$, if the buyer expects the next flash sale to take place at t_2, her reservation price $p_2(t; t_2)$ which makes her indifferent whether or not to accept satisfies

$$v_H - p_2(t; t_2) = \overbrace{\int_t^{t_2} (v_H - p_1(t')) e^{-\lambda(t'-t)} \lambda \mathrm{d}t'}^{\text{one buyer arrives in } [t, t_2]} + \overbrace{e^{-\lambda(t_2 - t)} \big(\beta(v_H - v_L) + (1-\beta)(v_H - p_1(t_2))\big)}^{\text{no buyer arrives in } [t, t_2]} . \qquad (10)$$

Combining equations (1) and (10), we can write

$$p_2(t;t_2) = v_H - \beta(v_H - v_L)\left(e^{-\lambda(1-t)}(1-\beta+\lambda(t_2-t)) + e^{-\lambda(t_2-t)}\right) \tag{11}$$

for all $t_2 \in [0,1]$ and $t \leq t_2$. Notice that $p_2(\cdot;t_2)$ has a bounded slope and is continuous in $t \in [0,t_2]$. In equilibrium, the buyer's belief is consistent with the seller's strategy, so $p_2(t;\tau_2(t))$ is the highest price that the buyer is willing to accept at time t both on and off the path of play.

Similarly, the payoff of the seller at t can also be written in terms of $\tau_2(t)$. Indeed, for each $t \in [0,1]$ and $t_2 \in [t_2,1]$ we define

$$\Pi_2(t;t_2) \equiv \overbrace{\int_t^{t_2}(p_2(t',t_2)+\Pi_1(t'))e^{-\lambda(t'-t)}\lambda \mathrm{d}t'}^{\text{one buyer arrives in } [t,t_2]} + \overbrace{e^{-\lambda(t_2-t)}(v_L+\Pi_1(t_2))}^{\text{no buyer arrives in } [t,t_2]}, \tag{12}$$

so the equilibrium payoff of the seller at state $(t,2)$ is $\Pi_2(t;\tau_2(t))$ regardless of his history.[23] The right-hand side of equation (12) can be interpreted as the payoff of the seller in state $(t,2)$ in a putative equilibrium where the next flash sale after t occurs at t_2. It is divided between the expected payoff of the seller if a buyer arrives at some time $t' \in [t,t_2]$, which is equal to $p_2(t';t_2)+\Pi_1(t')$, and his payoff if no buyer arrives in $[t,t_2]$. Therefore, the seller holds a flash sale at time t_2 and obtains $v_L + \Pi_1(t_2)$.

Step 2. *There is a unique equilibrium in which* $\tau_2(t) = \max\{t, t_2^*\}$ *for some* $t_2^* \in [0,1)$.

Proof. The proof proceeds as follows. First, we obtain a necessary condition for the existence of some $t \in [0,1)$ such that $\tau_2(t) > t$. This condition is written in terms of a function $\phi_2(\cdot)$ and the existence of some t_2 such that $\phi_2(t_2) = 0$. If $\phi_2(t_2) > 0$ for all t_2, it is necessarily the case that $\tau_2(t) = t$ for all t, and we verify that there is a unique equilibrium with this property. We then show that if $\phi_2(0) \leq 0$, then there is a unique $t_2^* \in [0,1)$ such that $\phi_2(t_2^*) = 0$. In this case, the unique equilibrium is such that $\tau_2(t) = \max\{t, t_2^*\}$.

Assume that there is an equilibrium and $t < 1$ such that $t_2 \equiv \tau_2(t) > t$. A necessary condition under which the seller has no incentive to hold a flash sale in $[t,t_2)$ is

$$\Pi_2(t';t_2) \geq \underbrace{\Pi_1(t') + v_L,}_{\text{holding flash sales immediately}} \qquad \forall t' \in [t,t_2).$$

[23]This is because in the continuation play, an accumulated buyer does not increases the seller's revenue in future regular sales (by Lemma 1 she never rechecks) or in future flash sales (all units will be sold instantaneously at price v_L).

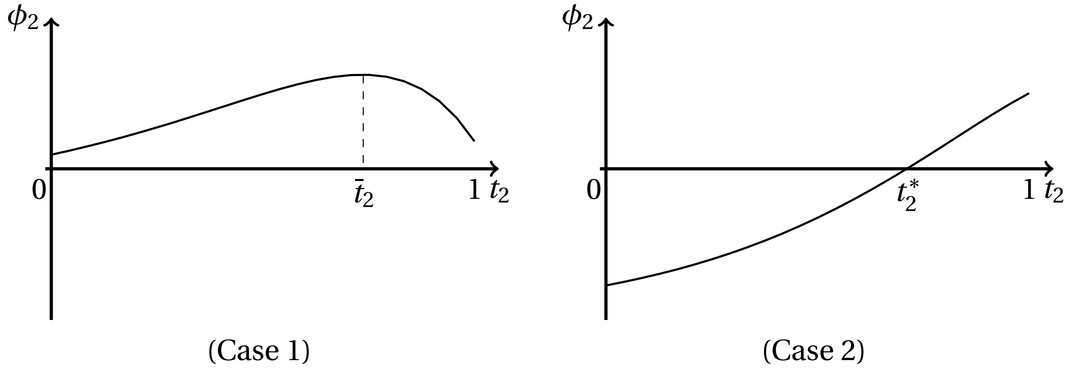

Figure 4: Case 1 occurs when $\phi_2(t_2) > 0$ for all $t_2 \in [0,1]$. In Case 2, $\phi_2(t_2) = 0$ at exactly one $t_2 \in (0,1)$, which defines t_2^* in this case.

Since $\Pi_2(t_2; t_2) = \Pi_1(t_2) + v_L$ (by equation (12)), this condition holds only if

$$0 \geq \phi_2(t_2)\ , \tag{13}$$

where

$$\phi_2(t_2) \equiv \lim_{t \nearrow t_2} \frac{\mathrm{d}(\Pi_2(t;t_2) - \Pi_1(t))}{\mathrm{d}t} = (v_H - v_L)\lambda\Big(\beta - 1 + e^{-(1-t_2)\lambda}(1-\beta^2+\beta(1-t_2)\lambda\Big)\ . \tag{14}$$

If equation (13) fails, then for some positive and small ε, we have

$$\Pi_2(t_2 - \varepsilon; t_2) < \Pi_1(t_2 - \varepsilon) + v_L\ .$$

Therefore, the seller has the incentive to deviate by holding a flash sale at $t_2 - \varepsilon$. Simple algebra shows that the derivative of the right-hand side of equation (14) with respect to t_2 (which is equal to $\phi_2'(t_2)$) equals 0 only at $\bar{t}_2 \equiv 1 - \frac{\beta^2+\beta-1}{\beta\lambda}$, and the second derivative at this point is negative. Hence, $\bar{t}_2$ is a local maximum of $\phi_2(\cdot)$. Note also since that $\phi_2(1) = (v_H - v_L)\lambda\beta(1-\beta) > 0$, there is no equilibrium where $\tau_2(t) = 1$ for some $t < 1$. Since $\bar{t}_2$ is the unique critical point of ϕ_2, there exists at most one value for $t_2 \in [0,1)$ that solves $\phi_2(t_2) = 0$. If a unique solution to $\phi_2(t_2) = 0$ exists, we denote it by t_2^*, and we will show that it has the property that $\tau_2(t) = \max\{t, t_2^*\}$. If, instead, $\phi_2(t_2) > 0$ for all $t_2 \in [0,1)$, we will show that it is the case that $t_2^* = 0$.[24] We study the two cases separately, which are presented in Figure 4.

- **Case 1 ($\phi_2(t_2) > 0$ for all $t_2 \in [0,1]$).** We claim that in the unique equilibrium, $\tau_2(t) = t$ for all

[24] Using expression (1) for p_1 and (3) for Π_1, we have that $\phi_2(t_2^*)/\lambda$ coincides with the left-hand side of equation (5) and thus characterizes the value of t_2^*, as argued in the main text.

$t \in [0,1]$; that is, the seller holds a flash sale immediately at any time t if $K_t = 2$. Therefore, $f_2(t) = 1$ for all $t \in [0,1)$, and $f_2(1) = 2$.

Assume by way of contradiction, that there is an equilibrium and a time $t \in [0,1)$ such that $\tau_2(t) > t$; that is, the seller does not hold flash sale immediately at state $(t,2)$. Therefore, it must be the case that the seller has no strict incentive to hold a flash sale in $[t, \tau_2(t))$. As we proved, this happens only if $\phi_2(\tau_2(t)) \leq 0$ (see equation (13)), which contradicts that $\phi_2(t_2) > 0$ for all $t_2 \in [0,1]$. Therefore, we have that if an equilibrium exists, it satisfies $\tau_2(t) = t$ for all t.

We now show the existence of an equilibrium satisfying that $\tau_2(t) = t$ for all $t \in [0,1]$. If such an equilibrium exists, the reservation price of a buyer arriving at time t and observing $K_t = 2$ is

$$p_2(t;t) = v_H - \beta(v_H - v_L) + (1-\beta)\beta e^{-(1-t)\lambda}(v_H - v_L) \,, \tag{15}$$

which is obtained by plugging the equilibrium condition $\tau_2(t) = t$ into (11). Fix a time $t \in [0,1)$. The equilibrium payoff of the seller is

$$\Pi_2(t;t) = \Pi_1(t) + v_L. \tag{16}$$

The payoff of the seller from waiting until $\hat{t}_2 > t$ to hold a flash sale while posting the equilibrium price in $(t, \hat{t}_2]$ is

$$\hat{\Pi}_2(t;\hat{t}_2) \equiv \int_t^{\hat{t}_2} (p_2(t', t') + \Pi_1(t'))e^{-\lambda(t'-t)}\lambda \mathrm{d}t' + e^{-\lambda(\hat{t}_2 - t)}(v_L + \Pi_1(\hat{t}_2)) \,. \tag{17}$$

Notice that the main difference between the previous equation and equation (12) is that now, in the integral term, we have $p_2(t', t')$ given by (15) instead of $p_2(t', t_2)$. This is because, by assumption, a buyer at each $t' \in [t, \hat{t}_2]$ expects the seller to hold a flash sale immediately, whereas the seller deviates and holds flash sale at $\hat{t}_2$.

We want to show that the seller wants to hold flash sale as soon as possible; i.e.,

$$t \in \arg\max_{\hat{t}_2 \in [t,1]} \hat{\Pi}_2(t;\hat{t}_2).$$

Given that $\hat{\Pi}_2(t;\cdot)$ is differentiable, we can differentiate it and obtain that

$$\frac{\partial \hat{\Pi}_2(t;\hat{t}_2)}{\partial \hat{t}_2} = e^{-\lambda(\hat{t}_2-t)}\phi_2(\hat{t}_2) < 0 \,. \tag{18}$$

Then,

$$\hat{\Pi}_2(t;\hat{t}_2) = \hat{\Pi}_2(t;t) + \int_t^{\hat{t}_2} \frac{\partial \hat{\Pi}_2(t;\hat{t}_2')}{\partial \hat{t}_2'} \mathrm{d}\hat{t}_2' < \hat{\Pi}_2(t;t) = \Pi_2(t;t) \,, \tag{19}$$

where the last equality holds by (16). Therefore, we conclude that the seller's payoff from this deviation is lower than his equilibrium payoff. Proceeding similarly, it is easy to see that if the seller charges a different price in $[t,\hat{t}_2]$, his payoff from the deviation is dominated by $\hat{\Pi}_2(t;\hat{t}_2)$. Indeed, charging a price strictly lower than $p_2(t';t')$ for some $t' \in [t,\hat{t}_2]$ is clearly suboptimal. Charging a price strictly above $p_2(t';t')$ for some $t' \in [t,\hat{t}_2]$ leads to no transaction at these times but increases the probability of transaction afterwards. The seller has no incentive to do so because the continuation payoff from a regular sale $p_2(t';t')+\Pi_1(t')$ is decreasing in t'.

The previous arguments show that there is an equilibrium satisfying that $\tau_2(t) = t$ for all $t \in [0,1]$, and such a condition uniquely determines the price offered by the seller as well as his flash-sale policy. Therefore, an equilibrium exists and is unique.

- **Case 2 ($\phi_2(t_2) = 0$ for a unique $t_2 \in [0,1]$).** Recall that the unique solution to $\phi_2(t_2) = 0$ in $[0,1]$ is denoted by t_2^*. Therefore, $\phi_2(t_2) < 0$ for all $t < t_2^*$ and $\phi_2(t_2) > 0$ for all $t_2 > t_2^*$. We claim that, in the unique equilibrium,

$$\tau_2(t) = \begin{cases} t_2^* & \text{if } t \le t_2^*, \\ t & \text{if } t > t_2^*, \end{cases}$$

or, $\tau_2(t) = \max\{t, t_2^*\}$; that is, $f_2(t) = 0$ for all $t < t_2^*$ and $f_2(t) = 1$ for all $t \in [t_2^*, 1)$. The same argument as in Case 1 can be used to show that in all equilibria, $\tau_2(t) = t$ for all $t \ge t_2^*$. It is left to verify that in all equilibria, $\tau_2(t) = t_2^*$ for all $t < t_2^*$; that is, the seller does not have the incentive to deviate.

Assume by way of contradiction, that there is an equilibrium and $t < t_2^*$ such that $\tau_2(t) < t_2^*$. Therefore, we have

$$\Pi_2(\tau_2(t);\tau_2(t)) = \Pi_1(\tau_2(t)) + v_L \,.$$

Consider the following deviation at $\tau_2(t)$ by the seller. Instead of holding a flash sale at time $\tau_2(t)$, the seller offers the regular price in $[\tau_2(t), t_2^*]$ (and holds no flash sales in $[\tau_2(t), t_2^*)$) and, if no unit has been sold before t_2^*, he holds a flash sale for one unit at t_2^*. By using equations (11) and (15), we can verify that for all $t' \in [t, t_2^*)$, we have that $p_2(t'; \tau_2(t')) \geq p_2(t'; t')$. This is intuitive: the reservation value of a buyer arriving at time t is maximized when she expects the first flash sale to be held immediately. Therefore, the payoff of the seller from the deviation is at least $\hat{\Pi}_2(\tau_2(t); t_2^*)$ as defined in equation (17). We then have

$$\Pi_1(\tau_2(t)) + v_L = \Pi_2(\tau_2(t), \tau_2(t)) \geq \hat{\Pi}_2(t; t_2) > \Pi_1(t) + v_L\,, \tag{20}$$

which contradicts the fact that $\Pi_1(\cdot)$ is decreasing. The weak inequality in equation (20) holds because the payoff from the deviation cannot be higher than the equilibrium payoff. The strict inequality in equation (20) is derived as follows. Simple algebra shows that for any $t_2 \in [0, t_2^*]$ and $t < t_2$, we have that

$$\frac{\partial}{\partial t}\left(e^{-\lambda t}\left(\hat{\Pi}_2(t; t_2^*) - \Pi_1(t) - v_L\right)\right) = e^{-\lambda t}\phi_2(t) < 0\,.$$

Since $e^{-\lambda t}(\hat{\Pi}_2(t; t_2) - \Pi_1(t) - v_L)$ is positive if and only if $\hat{\Pi}_2(t; t_2) - \Pi_1(t) - v_L$ is positive, we have

$$\hat{\Pi}_2(t; t_2) > \Pi_1(t) + v_L \ \text{ for all } t < t_2 \leq t_2^*. \tag{21}$$

Therefore, we conclude that in any equilibrium, $\tau_2(t) = t_2^*$ for all $t < t_2^*$ and, therefore, $\tau_2(t) = \max\{t, t_2^*\}$ for all $t \in [0, 1]$. In the rest of the proof, we show that the seller has no incentive to deviate.

First, assume that $t < t_2^*$ and consider a deviation consisting of holding a flash sale for one unit at time t. If the seller does so, he obtains $v_L + \Pi_1(t)$. Nevertheless, the payoff from following the equilibrium strategy and holding the flash sale at t_2^* (if there is no regular sale before then) is, by the previous argument, not lower than $\hat{\Pi}_2(t; t_2^*)$. Since, by equation (21), $\hat{\Pi}_2(t; t_2^*) > \Pi_1(t) + v_L$, we reach a contradiction. Hence, holding an early flash sale before t_2^* is suboptimal.

Second, assume again that $t < t_2^*$, and consider a deviation consisting of waiting to hold a flash sale for one unit at some time $t_2 > t_2^*$, while offering the equilibrium regular price $p_2(t')$ for all $t' \in [t, t_2]$. Notice that the seller's payoff from following the deviation strategy differs from his equilibrium payoff only if no buyer arrives in $[t, t_2^*]$. Therefore, the payoff from the deviation can be written as

$$\Pi_2(t; t_2^*) + e^{-\lambda(t_2^* - t)}(\hat{\Pi}_2(t_2^*; t_2) - \Pi_1(t_2^*) - v_L) , \tag{22}$$

where the first term is the seller's equilibrium continuation payoff, and the second term replaces the equilibrium payoff of the seller at t_2^* (equal to $\Pi_1(t_2^*) + v_L$) by his continuation payoff under the deviation (equal to $\hat{\Pi}_2(t_2^*; t_2)$). Proceeding similarly as we obtained equation (21), we have

$$\hat{\Pi}_2(t; t_2) < \Pi_1(t) + v_L \text{ for all } t_2^* \leq t < t_2.$$

where the inequality is the opposite of the one in (21), as now we use $\phi_2(t) > 0$ for all $t \in (t_2^*, 1]$. Therefore, we have that the second term in equation (22) is negative and hence the seller is worse off deviating. Consequently, again, this deviation is suboptimal for the seller.

Finally, assume that $t \geq t_2^*$, and consider a deviation consisting of waiting to hold a flash sale for one unit at some $t_2 > t$, while charging the equilibrium regular price $p_2(t')$ for $t' \in (t, t_2]$. In this case, proceeding as in Case 1, we have that equation (19) holds when $\hat{t}_2$ is replaced by t_2. As a result, the seller is better off not deviating.

The previous arguments show that the seller has no incentive to deviate from the flash-sale times while following the equilibrium regular price. By a similar argument to that of Case 1, we can see that the seller finds it suboptimal to deviate from the equilibrium regular price (equal to $p_2(t; \tau_2(t))$).

□

□

A.3 Proof of Results in Section 3.3

We begin this proof by first providing an additional result characterizing the equilibrium for the two-unit case. Let $\Pi_2(t)$ denote the seller's equilibrium continuation payoff at time t if $K_t = 2$, which equals $\Pi_2(t; t_2^*)$ if $t < t_2^*$, and $\Pi_1(t) + v_L$ if $t \geq t_2^*$ (recall that $\Pi_2(\cdot;\cdot)$ is defined in equation (12)). Notice that $\Pi_2(\cdot)$ is continuous. We show that $\Pi_2(\cdot)$ satisfies the so-called *smooth pasting condition.*

Lemma A1. *When* $K = 2$, *if* $t_2^* \in (0, 1)$, *then* $\lim_{t \nearrow t_2^*} \Pi_2'(t) = \lim_{t \searrow t_2^*} \Pi_2'(t) = \Pi_1'(t_2^*)$.

Proof. Notice that when $t_2^* \in (0,1)$, $\phi_2(t_2^*) = 0$ where $\phi_2(\cdot)$ is defined in (14). We obtain

$$\lim_{t \nearrow t_2^*} \frac{\mathrm{d}(\Pi_2(t;t_2^*) - \Pi_1(t))}{\mathrm{d}t} = \phi_2(t_2^*) = 0 \,.$$

Furthermore, when t approaches t_2^* from the right (and hence the profits at t when $K_t = 2$ are $\Pi_2(t;t)$, which is equal to $\Pi_1(t) + v_L$), we have that

$$\lim_{t \searrow t_2^*} \frac{\mathrm{d}(\Pi_2(t;t) - \Pi_1(t))}{\mathrm{d}t} = \lim_{t \searrow t_2^*} \frac{\mathrm{d}(v_L)}{\mathrm{d}t} = 0 \,.$$

Therefore, when $t_2^* \in (0,1)$, the smooth pasting condition holds: the marginal change in Π_2 and Π_1 around t_2^* is the same.

□

Proof of Proposition 3. The proof proceeds by induction. Fix $K \geq 3$. We assume that for all $k \leq K-1$ there exists a unique equilibrium as described in the statement of Proposition 3 (notice that, by Propositions 1 and 2, we know the result is true for $K = 1,2$). In this equilibrium, we use $\Pi_k(t)$ to denote the continuation payoff of the seller in state (t,k), for all $t \in [0,1]$ and $k = 1,\dots,K-1$. We want to prove that the result also applies to $k = K$. In particular, we assume that for all $2 \leq k \leq K-1$ there is some $t_k^* \in [0,1]$ satisfying $t_k^* < t_{k-1}^*$ if $t_{k-1}^* > 0$ and $t_k^* = 0$ if $t_{k-1}^* = 0$ such that:

- **Property 1:** $f_k(t) = \max\{k - \bar{k}, 0\}$ when $t \in [t_{\bar{k}}^*, t_{\bar{k}-1}^*)$ for some $1 \leq \bar{k} \leq k$.
- **Property 2:** If $t \in [t_{\bar{k}}^*, t_{\bar{k}-1}^*)$ for some $\bar{k}$, then, if $\bar{k} \geq k$, it is the case that $\Pi_k(t) > \Pi_{k-1}(t) + v_L$, while if $\bar{k} < k$, it is the case that $\Pi_k(t) = \Pi_{k-1}(t) + v_L$.
- **Property 3:** Π_k is differentiable and, if $t_k^* > 0$, then $\Pi_k'(t_k^*) = \Pi_{k-1}'(t_k^*)$.

Notice that, by Proposition 2 and Lemma A1, the above properties hold when $k = 2$. We will later prove that they hold for $2 \leq k \leq K$.

Consistently with the proof of Proposition 2, for each time $t \in [0,1]$ and $K_t = K$, $\tau_K(t)$ denotes the next time when a flash sale occurs if no regular sale takes place before $\tau_K(t)$ in a hypothetical equilibrium. That is,

$$\tau_K(t) \equiv \inf\{t' \geq t | f_K(t') > 0\} \in [t,1] \cup \{\infty\} \,.$$

Then, constructively, we prove that an equilibrium exists, that it is unique, and that it satisfies that there is some t_K^* such that $\tau_K(t) = t_K^*$ for $t < t_K^*$ and $\tau_K(t) = t$ for $t \geq t_K^*$. Finally, we prove that Properties 1, 2, and 3 hold for K.

Step 1. *In any equilibrium, the seller never keeps K units until the deadline,* $\tau_K(t) < 1$*, for all* $t < 1$.

Proof. Assume to the contrary that $\tau_K(t) = 1$ for some $t < 1$; then, $\tau_K(t') = 1$ for all $t' \in [t, 1]$. Therefore, assume without loss of generality that $t \in (t_2^*, 1)$. Assume also that there are $r \in \{1, ..., K-1\}$ regular sales at time t. Note that, by the induction hypothesis, for any time $t' > t$, we have $f_{K-r}(t') = K - r - 1$. Therefore, the seller holds a flash sale for all remaining units except for one after there is a regular sale at time t. The reservation value of a buyer at time $t \in (t_2^*, 1)$ is then $p_K(t; 1)$ which satisfies

$$\begin{aligned} v_H - p_K(t;1) = e^{-(1-t)\lambda}\beta_K(v_H - v_L) + e^{-(1-t)\lambda}(1-t)\lambda\big(\beta_{K-2} + (1-\beta_{K-2})\beta\big)(v_H - v_L) \\ + \big(1 - e^{-(1-t)\lambda} - e^{-(1-t)\lambda}(1-t)\lambda\big)\beta_{K-2}(v_H - v_L), \end{aligned}$$

where, as in the main text, we use β_k to denote the probability that an accumulated buyer obtains a good at a flash sale if k goods are offered and there is no other accumulated buyer, for each $k \le K$. Notice that if the buyer at time t rejects the offer and no other buyer arrives, she gets the good with probability β_K in the flash sale at the deadline. If, instead, exactly one buyer arrives in $[t, 1]$, say at time t', the t-buyer gets a good with probability β_{K-2} at the time-t' flash sale and, if she does not, she gets the remaining good with probability β in the flash sale at the deadline; that is, she gets a good with a total probability of $\beta_{K-2} + (1-\beta_{K-2})\beta = \beta_{K-1}$. Finally, if more than two buyers arrive (and the first of them arrives at time t'), the t-buyer gets a good only with probability β_{K-2} at the time-t' flash sale. The payoff of the seller at time t is given by

$$\Pi_K(t;1) \equiv \int_t^1 \big(p_K(t';1) + (K-2)v_L + \Pi_1(t')\big)e^{-\lambda(t'-t)}\lambda \mathrm{d}t' + e^{-\lambda(1-t)}Kv_L .$$

Then, proceeding analogously to the $K = 2$ case, we can differentiate $\Pi_K(t;1) - \Pi_1(t) - (K-1)v_L$ with respect to t and obtain that the derivative is positive at $t = 1$, a contradiction. □

Step 2. *Assume that* $K_t = K$ *at time* $t > t_{K-1}^*$*. Then, the number of units sold in the flash sale is* $f_K(t) = K - k$*, where k is such that* $t \in [t_{k+1}^*, t_k^*)$.

Proof. By way of contradiction, assume that there is an equilibrium where the seller does not hold a flash sale immediately at any time $t > t_{K-1}^*$ such that $K_t = K$ (if there is no regular sale at t); that is, $\tau_K(t) > t$. There are the following two possibilities:

- Assume first that $\tau_K(t) \in (t_{k+1}^*, t_k^*)$. We claim that the seller must offer $f_K(\tau_K(t)) = K - k$ units in the time-$\tau_K(t)$ flash sale. Note that if at time t the state is (t, K), then the seller is still in his private stage 0 and $F_{t^-}(h_0^s) = 0$, where $h_0^s = ((0,K),(0,0))$. In words, neither regular sales nor flash sales have taken place up to t. Suppose that $f_K(\tau_K(t)) < K-k$; that is, the remaining number of units is higher than k. By Property 1, the seller sells excess units above k immediately after t; that is, $f_{k'}(t') = k-k'$ for all $k' > k$ and $t' \in (t, t_k^*)$, which holds in particular for $k' = K - f_K(\tau_K(t))$. Hence, we have $F_t(h_0^s) = f_K(\tau_K(t))$ and, for all $t' > t$, $F_{t'}(h_0^s) = K-k$, which contradicts the assumption that $F_t(h_0^s)$ is a right continuous function of t. Hence, it must be that $f_K(\tau_K(t)) \in \{1,...,K-k\}$. Furthermore, notice that k is the unique maximizer of

$$(K-k')v_L + \Pi_{k'}(\tau_K(t)),$$

among all $k' \in \{1,...,K-k\}$, where the objective function represents the seller's continuation value from holding a k'-unit flash sale at time $\tau_K(t)$.

Now, the reservation price of a buyer arriving at state (t,K), with $t \in [t_{k+1}^*, \tau_K(t)]$, satisfies

$$\begin{aligned} v_H - p_K(t;\tau_K(t)) = &\int_t^{\tau_K(t)} \left(\beta_{K-k-1}(v_H - v_L) + (1-\beta_{K-k-1})(v_H - p_k(t'))\right) e^{-(t'-t)\lambda}\lambda \mathrm{d}t' \\ &+ e^{-(\tau_K(t)-t)\lambda}\left(\beta_{K-k}(v_H - v_L) + (1-\beta_{K-k})(v_H - p_k(\tau_K(t)))\right). \end{aligned} \tag{23}$$

The right-hand side of equation (23) represents the buyer's continuation value if she declines the offer. In the first term, another buyer arrives and purchases at some time $t' \in (t, \tau_K(t))$, and then the seller holds a flash sale for $K-k-1$ units. With probability β_{K-k-1}, the t-buyer gets a flash-sale offer and, with the complementary probability, she does not get a flash-sale offer and her continuation value becomes $v_H - p_k(t')$. In the second term, no other buyer arrives until $\tau_K(t)$, and the seller holds a $(K-k)$-unit flash sale. Similarly, the t-buyer gets the flash-sale offer with probability β_{K-k} and, with the complementary probability, she does get a good in the flash sale and obtains a continuation value of $v_H - p_k(\tau_K(t))$.

Notice that, by equation (23),

$$p_K(\tau_K(t);\tau_K(t)) = \beta_{K-k}v_L + (1-\beta_{K-k})p_k(\tau_K(t)).$$

Also, at any time $t \in (t_{k+1}^*, \tau_K(t))$, we have

$$p_{K-1}(\tau_K(t)) = \beta_{K-1-k}v_L + (1-\beta_{K-1-k})p_k(\tau_K(t)) > p_K(\tau_K(t);\tau_K(t)),$$

where the inequality holds because $\beta_{K-k-1} < \beta_{K-k}$ and $p_k(\tau_K(t)) > v_L$. Formalizing an analogous argument as at the end of Section 3.3 in the main text, the seller has an incentive to hold a flash sale earlier, i.e., at $\tau_K(t)-\varepsilon$, which is a contradiction.

- Assume now that $\tau_K(t)=t^*_{k+1}$. Since, by the induction hypothesis, $\Pi_{k+1}(t^*_{k+1})=\Pi_k(t^*_{k+1})+v_L$, we have that $(K-k')v_L+\Pi_{k'}(t^*_{k+1})$ is maximized with respect to k' at both $k+1$ and k. Consequently, either $f_K(t^*_{k+1})=K-k$ or $f_K(t^*_{k+1})=K-k-1$. If $f_K(t^*_{k+1})=K-k$ then the same argument as the one in the first case applies because

$$\begin{aligned} p_K(t^*_{k+1};t^*_{k+1}) &= \beta_{K-k}v_L+(1-\beta_{K-k})p_k(t^*_{k+1}) \\ &< \beta_{K-1-k}v_L+(1-\beta_{K-1-k})p_k(t^*_{k+1})=p_{K-1}(t^*_{k+1})\ . \end{aligned}$$

  If, instead, $f_K(t^*_{k+1})=K-k-1$, we have that

$$\begin{aligned} p_K(t^*_{k+1};t^*_{k+1}) &= \beta_{K-k-1}v_L+(1-\beta_{K-k-1})p_{k+1}(t^*_k) \\ &< \beta_{K-1-k}v_L+(1-\beta_{K-1-k)})p_k(t^*_{k+1})=p_{K-1}(t^*_{k+1})\ , \end{aligned}$$

  where the inequality holds because $p_{k+1}(t^*_{k+1})=\beta v_L+(1-\beta)p_k(t^*_{k+1})<p_k(t^*_{k+1})$. Therefore, the same argument as before applies.

□

Step 3. *If $t^*_{K-1}>0$ then $\tau_K(t)<t^*_{K-1}$ for all $t<t^*_{K-1}$.*

Proof. Assume by way of of contradiction that $t^*_{K-1}>0$ and $\tau_K(t)=t^*_{K-1}$ for some $t<t^*_{K-1}$ (notice that, by Step 2 of this proof, $\tau_K(t)\le t^*_{K-1}$ for all $t<t^*_{K-1}$). The argument is analogous to those used in Step 2, since now we have

$$\begin{aligned} p_K(t^*_{K-1};t^*_{K-1}) &= \beta_2 v_L+(1-\beta_2)p_{K-2}(t^*_{K-1}) \\ &< \beta v_L+(1-\beta)p_{K-2}(t^*_{K-1})=p_{K-1}(t^*_{K-1})\ . \end{aligned}$$

□

Step 4. *There is a unique equilibrium for $K_0=K$. In this equilibrium, $\tau_K(t)=\max\{t,t^*_K\}$ for some t^*_K such that $t^*_K\le t^*_{K-1}$, and $t^*_K<t^*_{K-1}$ if $t^*_{K-1}>0$.*

Proof. Note that, by the previous results, $\tau_K(t)=t$ for all $t\ge t^*_K$. We now proceed similarly to the proof in the case where $K=2$ (Step 2 in Proposition 2). Assume then that there is a time t such that $\tau_K(t)>t$. As we have shown in Steps 2 and 3 of this proof, it is the case that $\tau_K(t)>t$

only if $t < t^*_{K-1}$. Therefore, if the first flash sale after $t < t^*_{K-1}$ happens at $t_K = \tau_K(t) < t^*_{K-1}$, the equilibrium price satisfies

$$\begin{aligned} v_H - p_K(t;t_K) = &\int_t^{t_K} (v_H - p_{K-1}(t'))e^{-(t'-t)\lambda}\lambda \mathrm{d}t' \\ &+ e^{-(t_K-t)\lambda}\big(\beta(v_H - v_L) + (1-\beta)(v_H - p_{K-1}(t_K))\big) , \end{aligned} \tag{24}$$

and, similarly, the equilibrium payoff of the seller is

$$\begin{aligned} \Pi_K(t;t_K) = &\int_t^{t_K} (p_K(t';t_K) + \Pi_{K-1}(t'))e^{-\lambda(t'-t)}\lambda \mathrm{d}t' \\ &+ e^{-\lambda(t_K-t)}(v_L + \Pi_{K-1}(t_K)) . \end{aligned} \tag{25}$$

As in the two-unit case, assume that there is an equilibrium and time $t < 1$ such that $t < \tau_K(t)$, and let t_K denote $\tau_K(t)$. Note that the previous results imply that $t_K < t^*_{K-1}$ and, by (25), $\Pi_K(t_K;t_K) = \Pi_{K-1}(t_K) + v_L$. Thus, a necessary condition under which the seller has no incentive to hold a flash sale in $[t, t_K)$ is that $\Pi_K(t';t_K) \geq \Pi_{K-1}(t') + v_L$ for all $t' \in [t, t_K)$, which happens only if

$$\begin{aligned} 0 \geq \lim_{t' \nearrow t_K} \frac{\mathrm{d}(\Pi_K(t';t_K) - \Pi_{K-1}(t'))}{\mathrm{d}t'} &= -(1-\beta)\lambda\big(p_{K-1}(t_K) - v_L\big) - \Pi'_{K-1}(t_K) \\ &= \underbrace{\lambda\big((1-\beta)v_L - \Pi_{K-1}(t_K) + \Pi_{K-2}(t_K) + \beta p_{K-1}(t_K)\big)}_{\equiv \phi_K(t_K)} . \end{aligned}$$

Simple algebra shows that $\phi_K(t^*_{K-1}) = \lambda\beta(p_{K-1}(t^*_{K-1}) - v_L) > 0$. As in the proof of Proposition 2, we have two cases:

- **Case 1 ($\phi_K(t) > 0$ for all $t \in [0, t^*_{K-1})$).** Since $\phi_K(\tau_K(t)) \leq 0$ is a necessary condition for $\tau_K(t) > t$, there is no equilibrium where $\tau_K(t) > t$ for some $t \in [0,1]$ as in the two-unit case. To show the existence of an equilibrium, notice that if $\tau_K(t) = t$ is part of an equilibrium, the equilibrium price at time $t < t^*_{K-1}$ satisfies

$$v_H - p_K(t;t) = \beta(v_H - v_L) + (1-\beta)(v_H - p_{K-1}(t)) . \tag{26}$$

  Hence, the payoff of the seller from holding a flash sale at time $\hat{t}_K \in (t, t^*_{K-1})$ (instead of at time t) is bounded above by

$$\hat{\Pi}_K(t;\hat{t}_K) \equiv \int_t^{\hat{t}_K} (p_K(t';t') + \Pi_{K-1}(t'))e^{-\lambda(t'-t)}\lambda \mathrm{d}t' + e^{-\lambda(\hat{t}_K-t)}(v_L + \Pi_{K-1}(\hat{t}_K)) , \tag{27}$$

where, at each time $t' \in [t, \hat{t}_K]$, buyers expect the seller to hold a one-unit flash sale immediately, while the seller waits until $\hat{t}_K$ to do so and posts a regular price $p_K(t'; t')$. $\hat{\Pi}_K(t; \hat{t}_K)$ bounds the seller's payoff for the following reasons. If the seller, instead, charges a regular price different from $p_K(t'; t')$, his payoff is strictly lower, as in the two-unit case. Differentiating the previous expression with respect to $\hat{t}_K$ and evaluating $\hat{t}_K$ at t, we obtain an expression analogous to (18):

$$\frac{\partial \hat{\Pi}_K(t; \hat{t}_K)}{\partial \hat{t}_K}\bigg|_{\hat{t}_K = t} = -\phi_K(t) \; . \tag{28}$$

Then, arguing as in the proof of Proposition 2, we see that the seller has no incentive to delay the flash sale and therefore $\tau_K(t) = t$ for all t. Proceeding similarly, it is easy to see that if the seller charges a different price in $[t, \hat{t}_K]$, his payoff from the deviation is dominated by $\hat{\Pi}_K(t; \hat{t}_K)$. As a result, we have a unique equilibrium in this case.

- **Case 2 ($\phi_K(t) = 0$ for some $t \in [0, t^*_{K-1})$).** Now assume that $\phi_K(t) = 0$ for some $t \in [0, t^*_{K-1})$. Let $t^*_K < t^*_{K-1}$ be the maximum $t < t^*_{K-1}$ such that $\phi_K(t) = 0$ (notice that, by the continuity of ϕ_K, the maximum exists). Clearly, by the same argument as in Case 1 (and, in particular, equation (28)), in any equilibrium $\tau_K(t) = t$ for all $t \in [t^*_K, 1]$. The following result establishes that the seller does not have an incentive to hold a flash sale before t^*_K:

  Lemma A2. $\Pi_K(t; t^*_K) - \Pi_{K-1}(t) > v_L$ *for all* $t < t^*_K$.

  We leave the proof of Lemma A2 for the end of this proof. Notice that Lemma A2 leaves $\tau_K(t) = t^*_K$ for all $t < t^*_K$ and $\tau_K(t) = t$ for all $t \geq t^*_K$ as the unique candidate for an equilibrium. Indeed, if $\tau_K(t) < t^*_K$ for some $t < t^*_K$, we can take $\bar{t}$ to be the supremum of the times with this property. We have that the payoff of the seller at time $\bar{t}$, given by $v_L + \Pi_{K-1}(\bar{t})$, is lower than the payoff from deviating and waiting to hold a flash sale at time t^*_K, which is $\Pi_K(t; t^*_K)$. Hence, the seller has the incentive to hold a flash sale at t^*_K and not before then. Therefore, there is a unique equilibrium also in this case.

Given that, in the unique equilibrium, $\tau_K(t) = \max\{t, t^*_K\}$, we let $p_K(t)$ and

$$\Pi_K(t) \equiv \Pi_K(t; t^*_K)$$

denote the corresponding equilibrium values for the price and the profits in state (t, K). □

Step 5. *Properties 1, 2, and 3 hold for* K.

Proof. Property 1 holds for K because, as we showed in Steps 1 and 2, the seller holds a $(K-k)$-unit flash sales when $K_t = K$ and $t \in [t_k^*, t_{k-1}^*)$ for all $k \le K-1$. Property 2 holds for K because, by Step 4, $\tau_K(t) = t$ for all $t \ge t_K^*$ (so $\Pi_K(t) = \Pi_{K-1}(t) + v_L$), and $\Pi_K(t) > \Pi_{K-1}(t) + v_L$ for all $t < t_K^*$. Finally, Property 3 holds for K because of the definition of t_K^* (notice that t_K^* solves for $\phi_K(t_K^*) = 0$ whenever $t_K^* > 0$, which implies that $\Pi'_K(t_K^*) = \Pi'_{K-1}(t_K^*)$). □

It is left to prove Lemma A2.

Proof of Lemma A2. Define, for each $k \in \{1, ..., K\}$, $\Delta\Pi_k(t) \equiv \Pi_k(t) - \Pi_{k-1}(t)$, with the convention that $\Pi_0(t) = 0$ for all t. We want to prove that $\Delta\Pi_k(t) - v_L > 0$ for all $t < t_k^*$. To do this, notice that using some algebra we can obtain

$$p'_k(t) = \lambda(p_k(t) - p_{k-1}(t)) \quad \text{and} \tag{29}$$

$$0 = \Pi'_k(t) + \lambda\big(p_k(t) + \Pi_{k-1}(t) - \Pi_k(t)\big) \tag{30}$$

from equations (24) and (25) (using $\tau_K(t) = t_K^*$). Furthermore, using these equations and some algebra, it is easy to show that

$$\frac{\mathrm{d}}{\mathrm{d}t}\big(e^{-\lambda t}(\Delta\Pi_k(t) - v_L)\big) = -\lambda e^{-\lambda t}\big(\Delta\Pi_{k-1}(t) - \Delta p_k(t) - v_L\big)\,,$$

where it is convenient to define $\Delta p_k(t) = -(p_k(t) - p_{k-1}(t))$ to ensure that it is positive.

Note that $\Delta\Pi_k(t_k^*) - v_L = 0$. It is then only left to prove that the right-hand side of the previous expression is negative for $t < t_k^*$. The following lemma establishes the desired result (which corresponds to the case of $n = 1$):

Lemma A3. *For all $k = 2, ..., K$ and $n \le k$, there exists some $t_{k,n} \in [t_k^*, t_{k-1}^*]$ such that $\Delta\Pi_{k-1}(t) - n\Delta p_k(t) > v_L$ for all $t < t_{k,n}$ and $\Delta\Pi_{k-1}(t) - n\Delta p_k(t) < v_L$ for all $t \in (t_{k,n}, t_{k-1}^*]$.*

Proof. Consider first the case of $k = 2$. In this case, by Proposition 1, we have that, for $n = 0$, $\Delta\Pi_1(t) = \Pi_1(t)$ is decreasing for all t, and is equal to v_L only at $t_1^* = 1$. For $n = 1$ we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\big(e^{-\lambda t}(\Delta\Pi_1(t) - \Delta p_2(t) - v_L)\big) = \big(2\beta e^{-\lambda} - e^{-\lambda t}\big)\lambda(v_H - v_L)\,.$$

This is increasing in t. Also, the right-hand side of the expression is equal to 0 at $\bar{t} \equiv 1 - \frac{\log(2\beta)}{\lambda}$. It is only left to show that $t_2^* \le \bar{t}$. If $\beta \le \frac{1}{2}$, then the result is clear. Thus, assume $\beta > \frac{1}{2}$. In this case, notice that t_2^* is defined by $\phi_2(t_2^*) = 0$ (defined in (14)), where $\phi_2(t)$ is negative for all $t < t_2^*$. We then have

$$\phi_2(\bar{t}) = (v_H - v_L)\lambda\frac{1}{2}\left(\beta + \frac{1}{\beta} - 2 + \log(2\beta)\right) .$$

Notice that since $\beta > \frac{1}{2}$ we have that both $\beta + \frac{1}{\beta} > 2$ and $\log(2\beta) > 0$. Therefore, $\phi_2(\bar{t}) > 0$ and hence $\bar{t} > t_2^*$, which proves the result.

Now we proceed by induction: take $k > 2$ and assume that for all $\tilde{k} < k$ and for all $n \in \{0, ..., \tilde{k} - 1\}$, $\Delta\Pi_{\tilde{k}-1}(t) - n\Delta p_{\tilde{k}}(t)$ satisfies the property of the statement of the lemma for some $t_{\tilde{k},n}$. Simple algebra (again from equations (29) and (30)) shows that the following formula is true for all $k \geq 3$ and $n \in \{0, ..., k-1\}$:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(e^{-\lambda t}(\Delta\Pi_{k-1} - n\Delta p_k(t))\right) = -\lambda e^{-\lambda t}\left(\Delta\Pi_{k-2}(t) - (n+1)\Delta p_{k-1}(t)\right) . \tag{31}$$

We first want to show that the statement of the lemma holds for k and $n = 0$. In this case, note first that

$$\frac{\mathrm{d}}{\mathrm{d}t}(e^{-\lambda t}\Delta\Pi_{k-1}(t))\Big|_{t=t_{k-1}^*} = -\lambda e^{-\lambda t_{k-1}^*}\Delta\Pi_{k-1}(t_{k-1}^*) + e^{-\lambda t_{k-1}^*}\Delta\Pi'_{k-1}(t_{k-1}^*) = -e^{-\lambda t_{k-1}^*}v_L$$

where we used the fact that the value-matching and smooth-pasting conditions hold for $\Pi_{k-1}(t)$ at $t = t_{k-1}^*$ (Properties 2 and 3). Also, we evaluate the expression in (31) at $n = 0$:

$$\frac{\mathrm{d}}{\mathrm{d}t}(e^{-\lambda t}\Delta\Pi_{k-1}(t)) = -\lambda e^{-\lambda t}(\Delta\Pi_{k-2}(t) - \Delta p_{k-1}(t)) .$$

By assumption, there is at most one time $t \in [t_{k-1}^*, t_{k-2}^*]$ at which the right-hand side of the previous equation equals $-\lambda e^{-\lambda t}v_L$. Denote it by $t_{k-1,1}$. Due to its minus sign, the right-hand side of the previous equation is strictly below $-\lambda e^{-\lambda t}v_L$ for all $t < t_{k-1,1}$. Then, $\frac{d}{dt}(e^{-\lambda t}\Delta\Pi_{k-1}(t))$ is negative for all $t < t_{k-1}^*$, and so the result holds for $n = 0$.

Notice now that for all $n \in \{1, ..., k-1\}$ we have, again from equations (29) and (30), that

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(e^{-\lambda t}(\Delta\Pi_{k-1}(t) - n\Delta p_k(t))\right) = -\lambda e^{-\lambda t}(\Delta\Pi_{k-2}(t) - (n+1)\Delta p_{k-1}(t)) .$$

By the induction argument, again, the right-hand side of the previous equation equals to $-\lambda e^{-\lambda t}v_L$ in $[0, t_{k-2}^*]$ in at most one time, $t = t_{k-1,n+1}$, and due to the minus sign, it is strictly below $-\lambda e^{-\lambda t}v_L <$

0 for all $t < t_{k-1,n+1}$. Now, given that $\Delta p_{k-1}(t) \geq 0$ for all $t < t^*_{k-1}$, we have

$$\Delta\Pi_{k-2}(t) - (n+1)\Delta p_{k-1}(t) \leq \Delta\Pi_{k-2}(t) - \Delta p_{k-1}(t) < 0\,,$$

where the strict inequality follows from the case where $n = 0$. So, we have that the derivative of $e^{-\lambda t}(\Delta\Pi_{k-1}(t) - n\Delta p_k(t)))$ is negative for all $t < t^*_{k-1}$, and then the result holds. □

□

□

A.4 Proof of Results in Sections 3.4 and 4

Proof of Proposition 4. To examine the effect of a change in λ on the equilibrium outcome, we label the equilibrium regular price, the seller's profit, and the equilibrium flash-sale times as $p^\lambda_k(t)$, $\Pi^\lambda_k(t)$, and t^λ_k for $k = 1,...,K$, respectively.

Suppose that $\lambda' < \lambda$. We want to show that the corresponding flash-sale times satisfy, for all $k = 1,...,K$,

$$t^{\lambda'}_k \leq t^\lambda_k.$$

That is, for each k, the flash-sale time in the λ'-model is earlier than the one in the λ-model.

Consider an auxiliary model where we lower the λ-model's deadline from 1 to $\lambda'/\lambda < 1$ while keeping the buyer's arrival rate equal to λ. Since the equilibrium is solved backwardly, for each state (t,k) with $t \geq 0$, the continuation play in the auxiliary game is identical to that in the original λ model at state $(t + 1 - \lambda'/\lambda, k)$. Therefore, in the auxiliary game, for $t \in [0, \lambda'/\lambda]$ and $k = 1,...,K$, the equilibrium regular price and the seller's continuation value satisfy

$$p^a_k(t) = p^\lambda_k(t + 1 - \lambda'/\lambda)$$

and

$$\Pi^a_k(t) = \Pi^\lambda_k(t + 1 - \lambda'/\lambda),$$

respectively, and for each $k = 1,...,K$, the flash sale time satisfies

$$t^a_k = \max\{0, t^\lambda_k - 1 + \lambda'/\lambda\}.$$

Now consider the λ'-model. An induction argument shows that

$$p_k^{\lambda'}(t) = p_k^a\big(t\tfrac{\lambda'}{\lambda}\big) \text{ and } \Pi_k^{\lambda'}(t) = \Pi_k^a\big(t\tfrac{\lambda'}{\lambda}\big),$$

for each $k = 1, ..., K$, $t \in [0,1]$ in equilibrium. Additionally, if there is some $t \in [0,1)$ such that

$$\phi_k^a\big(t\tfrac{\lambda'}{\lambda}\big) \equiv \lambda\big((1-\beta)v_L - \Pi_{k-1}^a\big(t\tfrac{\lambda'}{\lambda}\big) + \Pi_{k-2}^a\big(t\tfrac{\lambda'}{\lambda}\big) + \beta p_{k-1}^a\big(t\tfrac{\lambda'}{\lambda}\big)\big) = 0,$$

then

$$\phi_k^{\lambda'}(t) \equiv \lambda((1-\beta)v_L - \Pi_{k-1}^{\lambda'}(t) + \Pi_{k-2}^{\lambda'}(t) + \beta p_{k-1}^{\lambda'}(t)) = 0,$$

and so

$$t_k^{\lambda'} = t_k^a \tfrac{\lambda}{\lambda'}.$$

Notice that when $\phi_k^a(\cdot) > 0$, $\phi_k^{\lambda'}(\cdot) > 0$ for all t, and $t_k^{\lambda'} = t_k^a = 0$ as well. It is then easy to see that

$$t_k^{\lambda} - t_k^{\lambda'} = \min\{t_k^{\lambda}, \big(\tfrac{\lambda}{\lambda'} - 1\big)(1 - t_k^{\lambda})\} \geq 0, \forall k,$$

and that the inequality is strict if $t_k^{\lambda} > 0$.

□

Proof of Proposition 5. Recall that t_2^* equals the unique solution to $\phi_2(t_2) = 0$, defined in equation (14), for $t_2 \in [0,1]$ if it exists; $t_2^* = 0$ otherwise. We want to prove that when the solution to $\phi_2(t_2) = 0$, denoted by $t_2(\beta)$, exists, it is decreasing in β. We first assume that β is such that $\phi_2(0) < 0$. Since $\phi_2(1) > 0$ for all values of $\beta \in (0,1)$ (see the proof of Proposition 2), we have $t_2(\beta) \in (0,1)$. We then differentiate equation (14) with respect to β. After some algebra, we obtain

$$\frac{\mathrm{d}t_2(\beta)}{\mathrm{d}\beta} = -\frac{2}{\lambda}(1 + e^{\lambda(1-t_2)})^{-1}\left(1 - \frac{1}{\sqrt{1 + 4\frac{1+e^{\lambda(1-t_2)}(1+\lambda(1-t_2)e^{\lambda(1-t_2)}(\lambda(1-t_2)-1))}{(\lambda(1-t_2)+e^{(1-t_2)\lambda}(e^{\lambda(1-t_2)})+\lambda(1-t_2))^2}}}\right)^{-1},$$

where, to ease notation, we omitted the explicit dependence of t_2 on β on the right-hand side of the expression. Notice that the term $(1 + \lambda(1-t_2)e^{\lambda(1-t_2)}(\lambda(1-t_2) - 1))$ in the numerator of the term inside the square root is weakly positive for all $t_2 \leq 1$, since it is equal to 0 when $t_2 = 1$ and its derivative with respect to t_2 is $-\lambda(1 + e^{\lambda(1-t_2)})\lambda(1-t_2) < 0$. Therefore, the right-hand side of the previous expression is negative, and hence $t_2(\cdot)$ is strictly decreasing. When, instead, β is such that $\phi_2(0) > 0$, we have that small changes in β do not affect t_2^*, which is equal to 0. The previous result implies that t_2^* is a decreasing function of β: it is strictly decreasing when $t_2^* > 0$ and weakly decreasing otherwise. Therefore, there is some value $\bar{\beta} \in [0,1]$ such that $t_2(\beta) > 0$ when $\beta < \bar{\beta}$ and $t_2(\beta) = 0$ when $\beta > \bar{\beta}$.

We finally prove that $\bar{\beta} \in (0,1)$. We first show that $\bar{\beta} < 1$. Note that $e^{-\lambda t}\phi_2(t)$ is a concave function:

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}(e^{-\lambda t}\phi_2(t)) = -(v_H - v_L)\lambda^3(1-\beta)e^{-\lambda t} < 0\,.$$

Also, $e^{-\lambda}\phi_2(1) = (v_H - v_L)\lambda(1-\beta)\beta e^{-\lambda} > 0$, and $\lim_{\beta\to 1}\phi_2(0) = (v_H - v_L)\lambda^2 e^{-\lambda} > 0$. So, by continuity of ϕ_2, if β is close to 1, both $\phi_2(0)$ and $\phi_2(1)$ are strictly positive and, since $e^{-\lambda t}\phi_2(t)$ is concave, $\phi_2(t)$ is also strictly positive for all $t \in [0,1]$, so $t_2^* = 0$ in the unique equilibrium. Second, we show that $\bar{\beta} > 0$. Suppose not, then for every $\beta \in (0,1)$, $t_2(\beta) = 0$. In other words, $\phi_2(t) \geq 0$ for all t. But

$$\phi_2(0) = \beta - 1 + e^{-\lambda}(1 - \beta^2 + \beta\lambda),$$

which is strictly negative when β is close to 0, a contradiction. □

Proof of Proposition 6. Using equation (1) it is clear that the result is true for $K = 1$. Then, we can proceed analogously to show $\lim_{\beta\to 0} t_2^* = 1$ and $\lim_{\beta\to 0} p_2(t) = v_H$ for all $t \in [0,1]$. The reason is that, when β is low, even when the time is close to the flash sale, the willingness to pay of the buyers is close to v_H, since obtaining the good in a flash sale is unlikely. As a result the seller has the incentive to postpone flash sales toward the deadline. This argument can be applied iteratively for any K. □

Proof of Proposition 7. First, consider the case where $K_0 = 1$, and for notational convenience, let $G(v_H) = \frac{v_H - \underline{v_H}}{\overline{v_H} - \underline{v_H}}$ denote the CDF of a uniform distribution in $[\underline{v_H}, \overline{v_H}]$. Assume that there is an equilibrium as described in Proposition 1, where now $p_1(t)$ is the price corresponding to the case where the valuation of the buyers is $\underline{v_H}$. Fix a rechecking cost c, if $\overline{v_H}$ is sufficiently close to $\underline{v_H}$, no type of buyer rechecks the price. Therefore, if a buyer with valuation v_H arrives at time t, the price she is willing to pay equals

$$p_1(t, v_H) \equiv v_H - e^{-\lambda(1-t)}(v_H - v_L)\beta > \underline{v_H} - e^{-\lambda(1-t)}(\underline{v_H} - v_L)\beta = p_1(t)\,.$$

Consider the seller's gain from deviating. By increasing the price from $p_1(t)$ to $p_1(t) + \varepsilon$, for some small $\varepsilon > 0$, the seller increases the price conditional on acceptance, but lowers the probability of acceptance conditional on arrival to $1 - G(v_H(t, p_1(t) + \varepsilon))$, where $v_H(t, p)$ is defined as follows:

$$p = v_H(t,p) - e^{-\lambda(1-t)}(v_H(t,p) - v_L)\beta \;\Rightarrow\; v_H(t,p) = v_L + \frac{p - v_L}{1 - e^{-(1-t)\lambda}\beta}\,.$$

Consider the benefit from deviating at all $s \in [t, t+\Delta]$ to offering $p_1(s) + \varepsilon$, for $\Delta > 0$ small. This deviation is not profitable only if $p_1(t)$ is higher than $(p_1(t)+\varepsilon)(1 - G(v_H(t, p_1(t)+\varepsilon)))$ for all $\varepsilon > 0$. The derivative of this last expression with respect to ε yields

$$1 - G(v_H(t, p_1(t)+\varepsilon)) - \frac{(p_1(t)+\varepsilon)G'(v_H(t,p_1(t)+\varepsilon))}{1-e^{(1-t)\lambda}\beta} < 1 - \frac{v_L}{\overline{v_H} - \underline{v_H}}\,.$$

If $\overline{v_H} - \underline{v_H}$ is small enough, the last term to the right of the previous function is negative, so it is optimal to choose $\varepsilon = 0$.

For a general K, a similar argument applies. Indeed, at any given time, the reservation price of a buyer with valuation is $v_H - b_{K_t}(t)(v_H - v_L)$, where $b_k(t)$ is the probability of obtaining a good in the future (in a flash sale) if at time t the stock is k, for all $k = 1, \ldots, K$. Therefore, in general, the payoff gain from offering a price equal to $p_1(t) + \varepsilon$ divided by ε, as $\varepsilon \to 0$,is

$$1 - G(v_H(t, p_1(t)+\varepsilon)) - \frac{(p_1(t)+\varepsilon)G'(v_H(t,p_1(t)+\varepsilon))}{1-b_{K_t}(t)} + O(\varepsilon) < 1 - \frac{v_L}{\overline{v_H} - \underline{v_H}}\,.$$

Therefore, again, if $\overline{v_H} - \underline{v_H}$ is sufficiently small, the right-hand side of the inequality of the previous function is negative, and so it is optimal to choose $\varepsilon = 0$. □

References

Akcigit, U. and Q. Liu (2015). The role of information in innovation and competition. *Journal of the European Economic Association 14*(4), 828–870.

Anton, J. J., G. Biglaiser, and N. Vettas (2014). Dynamic price competition with capacity constraints and a strategic buyer. *International Economic Review 55*(3), 943–958.

Aryal, G., C. Murry, and J. W. Williams (2018). Price discrimination in international airline markets. *Unpublished Manuscript, University of North Carolina Chapel Hill.*

Aumann, R. J. and B. Peleg (1974). A note on gale's example. *Journal of Mathematical Economics 1*(2), 209–211.

Bergin, J. and W. B. MacLeod (1993). Continuous time repeated games. *International Economic Review 34*(1), 21–37.

Biyalogorsky, E., E. Gerstner, and B. Libai (2001). Customer referral management: Optimal reward programs. *Marketing Science 20*(1), 82–95.

Board, S. (2008). Durable-goods monopoly with varying demand. *The Review of Economic Studies 75*(2), 391–413.

Board, S. and A. Skrzypacz (2016). Revenue management with forward-looking buyers. *Journal of Political Economy 124*(4), 1046–1087.

Chen, C.-H. (2012). Name your own price at priceline. com: Strategic bidding and lockout periods. *The Review of Economic Studies 79*(4), 1341–1369.

Conlisk, J., E. Gerstner, and J. Sobel (1984). Cyclic pricing by a durable goods monopolist. *The Quarterly Journal of Economics 99*(3), 489–505.

Dana, J. D. (1999). Equilibrium price dispersion under demand uncertainty: The roles of costly capacity and market structure. *The RAND Journal of Economics 30*(4), 632–660.

Dana, J. D. and E. Orlov (2014). Internet penetration and capacity utilization in the us airline industry. *American Economic Journal: Microeconomics 6*(4), 106–137.

Dana, J. D. and K. Williams (2018). Oligopoly price discrimination: The role of inventory controls. *Unpublished Manuscript, Yale University.*

Dana, Jr, J. D. (1998). Advance-purchase discounts and price discrimination in competitive markets. *Journal of Political Economy 106*(2), 395–422.

Deb, R. and M. Said (2015). Dynamic screening with limited commitment. *Journal of Economic Theory 159*, 891–928.

Dilme, F. and F. Li (2018). Revenue management without commitment: dynamic pricing and periodic fire sales. *Unpublished manuscript, University of Bonn and University of North Carolina Chapel Hill*.

Eden, B. (1990). Marginal cost pricing when spot markets are complete. *Journal of Political Economy 98*(6), 1293–1306.

Ely, J. C., D. F. Garrett, and T. Hinnosaar (2017). Overbooking. *Journal of the European Economic Association 15*(6), 1258–1301.

Escobari, D. (2012). Dynamic pricing, advance sales and aggregate demand learning in airlines. *The Journal of Industrial Economics 60*(4), 697–724.

Fuchs, W. and A. Skrzypacz (2013). Bargaining with deadlines and private information. *American Economic Journal: Microeconomics 5*(4), 219–243.

Gale, D. (1974). Exchange equilibrium and coalitions: An example. *Journal of Mathematical Economics 1*(1), 63–66.

Gale, I. L. and T. J. Holmes (1993). Advance-purchase discounts and monopoly allocation of capacity. *The American Economic Review 83*(1), 135–146.

Gallego, G. and G. Van Ryzin (1994). Optimal dynamic pricing of inventories with stochastic demand over finite horizons. *Management Science 40*(8), 999–1020.

Garrett, D. F. (2016). Intertemporal price discrimination: Dynamic arrivals and changing values. *American Economic Review 106*(11), 3275–99.

Gershkov, A. and B. Moldovanu (2009). Dynamic revenue maximization with heterogeneous objects: A mechanism design approach. *American Economic Journal: Microeconomics 1*(2), 168–198.

Gershkov, A., B. Moldovanu, and P. Strack (2017). Revenue-maximizing mechanisms with strategic customers and unknown, markovian demand. *Management Science 64*(5), 2031–2046.

Gul, F., H. Sonnenschein, and R. Wilson (1986). Foundations of dynamic monopoly and the coase conjecture. *Journal of Economic Theory 39*(1), 155–190.

Hart, O. and J. Tirole (1990). Vertical integration and market foreclosure. *Brookings papers on economic activity. Microeconomics 1990*, 205–286.

Horn, H. and A. Wolinsky (1988). Bilateral monopolies and incentives for merger. *The RAND Journal of Economics*, 408–419.

Hörner, J. and L. Samuelson (2011). Managing strategic buyers. *Journal of Political Economy 119*(3), 379–425.

Inderst, R. (2008). Dynamic bilateral bargaining under private information with a sequence of potential buyers. *Review of Economic Dynamics 11*(1), 220–236.

Kreps, D. M. and J. A. Scheinkman (1983). Quantity precommitment and bertrand competition yield cournot outcomes. *The Bell Journal of Economics 14*(2), 326–337.

Liu, Q., K. Mierendorff, X. Shi, and W. Zhong (2018). Auctions with limited commitment. *American Economic Review forthcoming*.

McAfee, R. P. and M. Schwartz (1994). Opportunism in multilateral vertical contracting: Nondiscrimination, exclusivity, and uniformity. *The American Economic Review*, 210–230.

McAfee, R. P. and V. Te Velde (2006). Dynamic pricing in the airline industry. *Handbook on Economics and Information Systems, Ed: TJ Hendershott, Elsevier*.

McAfee, R. P. and D. Vincent (1997). Sequentially optimal auctions. *Games and Economic Behavior 18*(2), 246–276.

McAfee, R. P. and T. Wiseman (2008). Capacity choice counters the coase conjecture. *The Review of Economic Studies 75*(1), 317–331.

Mierendorff, K. (2016). Optimal dynamic mechanism design with deadlines. *Journal of Economic Theory 161*, 190–222.

Murto, P. and J. Välimäki (2013). Delay and information aggregation in stopping games with private information. *Journal of Economic Theory 148*(6), 2404–2435.

Nocke, V. and M. Peitz (2007). A theory of clearance sales. *The Economic Journal 117*(522), 964–990.

Öry, A. (2017). Consumers on a leash: advertised sales and intertemporal price discrimination. *Unpublished manuscript, Yale University*.

Pai, M. M. and R. Vohra (2013). Optimal dynamic auctions and simple index rules. *Mathematics of Operations Research 38*(4), 682–697.

Sobel, J. (1991). Durable goods monopoly with entry of new consumers. *Econometrica: Journal of the Econometric Society 59*(5), 1455–1485.

Talluri, K. T. and G. J. Van Ryzin (2006). *The Theory and Practice of Revenue Management*, Volume 68. Springer Science & Business Media.

Williams, K. R. (2018). Dynamic airline pricing and seat availability. *Cowles Foundation Discussion Paper*.