# Nonlinear Pricing of Storable Goods[†]

*By* Igal Hendel, Alessandro Lizzeri, and Nikita Roketskiy[*]

*This paper develops a model of nonlinear pricing of storable goods. We show that storability imposes novel constraints on a monopolist's ability to extract surplus. We then show that the attempt to relax these constraints can generate cyclical patterns in pricing and sales, even when consumers are homogeneous. Thus, the model provides a novel explanation for sales that does not rely on discriminating heterogeneous consumers. Enriching the model to allow for buyer heterogeneity in storage technology delivers the prediction that larger bundles are more likely to be on sale.* (JEL D11, D42, L12)

Nonlinear pricing is prevalent in many markets, from phone and electricity tariffs to supermarkets items. There is an extensive literature that studies nonlinear pricing as a tool for surplus extraction, often as a device for price discrimination in the context of heterogeneous buyers (see Wilson 1997). However, the contrast with linear pricing is particularly stark when consumers are homogeneous. In this case, the optimal nonlinear pricing policy involves a monopolist selling the socially optimal quantity and extracting all the surplus.

Many of the products sold through nonlinear prices are storable. For example, the typical scanner data show quantity discounts in a variety of products ranging from yogurt to detergent (see Hendel and Nevo 2006a,b). Many other products, like intermediate goods, are also storable and priced nonlinearly. We say a good is storable if consumers can set aside units for later consumption. Product storability enables consumers to detach the timing of purchase from the timing of consumption. Storability has distinct implications from those well-studied in the durable good literature. While durable good purchases can also be timed, the literature on durable goods monopoly focuses on the case where consumers have unit demands.[1]

The goal of the present paper is to study how consumers' ability to store a product may affect sellers' abilities to extract surplus via nonlinear prices. For this purpose, it is important to allow consumers to demand multiple units.

[1] The central result in this literature is the Coase conjecture: if the discount factor is high, the equilibrium involves nearly efficient trading at prices close to the marginal cost. See (Waldman 2003) for a survey.

We study seller behavior in two related environments. We start our analysis with homogenous buyers so that there is no gain from screening consumers, and therefore the standard discrimination motive is absent. We then consider heterogeneity in ability to store. The analysis is presented through a combination of a two-period model and a particularly stark infinite horizon setup. The two-period model is a natural first step, since it helps to highlight some key forces and it is simple enough that a number of extensions can potentially be studied. However, the two-period framework introduces asymmetries between periods in that the second period involves no further potential for storage. The infinite horizon model restores symmetry across periods, but it is a more complicated environment, and we restrict the analysis to a subclass of the possible policies by the monopolist.

Storability introduces a dynamic form of participation constraints because the monopolist must ensure that the consumer does not wish to store in anticipation of skipping a purchase. We call these constraints "no-skipping" constraints. We show that the impact of these constraints can be severe. For instance, if the seller chooses a stationary policy of offering the same bundle over time, the monopolist can lose all ability to price nonlinearly. The monopolist is still able to make a profit, but, with such policies, nonlinear prices do no better than linear prices: by choosing the frequency of purchases and consuming out of storage, consumers fully undo any attempt to extract additional surplus. Storage enables consumers to unbundle nonlinear pricing policies. The logic of this result is related to the constraint on nonlinear pricing that is imposed by resale: the consumer can purchase marginal units cheaply in a bundle and "resell" them to his future self.

Given the ineffectiveness of such stationary policies, we investigate whether there are more sophisticated ways for the monopolist to enable surplus extraction via nonlinear prices. We show that the monopolist can partially regain some ability to extract surplus via a suitable cyclical policy that involves the infrequent sale of large bundles.[2] The main idea behind the value of cyclical policies is that the seller gains from forcing consumers to use their storage capacity as a way of relaxing the "no-skipping" constraint. By offering infrequent purchase opportunities, interpreted as sales, the seller limits consumers' opportunities to skip purchases and consume out of storage. At the same time, by making each purchase take up the full storage capacity, the seller makes the storage unavailable for skipping, thereby relaxing the no-skipping constraints. In other words, the infrequent sale of large quantities eliminates consumers' ability to get ready to skip, namely, to slowly accumulate inventory in anticipation of skipping a purchase. Notice that the cost of skipping is the utility of foregone consumption, and this cost goes down if consumers can save over many periods, so that the foregone consumption involves a lower utility loss.

While the quantity sold is determined by the storage capacity, consumption and prices are determined by the frequency of purchases. More frequent purchases of a given quantity translate into a higher consumption rate. Prices are determined

---

[2]Note that in our model this holds even in an environment with identical consumers. We interpret these cyclical policies are useful solely because they serve to enhance nonlinear prices, thereby presenting a novel reason for cyclical pricing (or sales) by a monopolist, and provide a stark contrast with models of sales based on a motive to discriminate among heterogeneous consumers, as in Salop and Stiglitz (1982); Narasimhan (1988); Sobel (1984); Hong, McAfee, and Nayyar (2002); and Pesendorfer (2002).

by the following trade-off. In order to make the consumer show up to purchase (namely, prevent skipping), the monopolist must promise him enough surplus so that he does not prefer to consume out of storage. It turns out that sellers can extract at most $2[V(c) - V(c/2)]$ (per period). The term $V(c)$ is the surplus at the intended consumption, while $V(c/2)$ is the surplus of a consumer who skips every second purchase.[3] A higher consumption rate (up to the socially efficient level) increases the target surplus, thus, the amount the seller can potentially capture. On the other hand, a higher consumption flow also increases the utility from skipping a purchase event. The enhanced threat of skipping limits the surplus that can be extracted. We show that consumption is distorted downward, relative to the efficient level (absent storage), but it is larger than consumption under optimal linear prices.

We extend the analysis to introduce heterogeneity in consumers' ability to store, a type of heterogeneity that naturally cannot arise in static models. This heterogeneity allows us to generate more realistic patterns of sales and to generate a novel empirical implication: the model predicts that sales are more important for large packages. While we emphasize that the model is very stark and does not allow for many features that are present in actual markets, it is interesting to note that this prediction is consistent with typical scanner data patterns. For example, Hendel and Nevo (2006b) report that while the small (32 oz.) detergent container is hardly on sale (2 percent of the time), the larger, most purchased, container size (128 oz.) is on sale 16.6 percent of the time. In the soda category, 12 and 24-packs are on sale twice as often as single soda cans (over 40 percent of the time versus 19.6 percent), while 6-packs are on sale 34.3 percent of the time.

## I. Related Literature

There is an extensive literature on nonlinear pricing (see, for instance, Wilson 1997 and Rochet and Stole 2002). The literature has considered many constraints on nonlinear pricing, including resale and asymmetric information on consumers' valuations, as well as competition among producers (see for instance Stole 2007). However, all the theory of nonlinear pricing is static, and ignores potential effects that arise from intertemporal substitution in demand.

Several theoretical papers offer models of price dispersion, (Varian 1980, Salop and Stiglitz 1982, Narasimhan 1988), interpreted as sales, however, these models do not capture the dynamics of demand generated by sales. Hong, McAfee, and Nayyar (2002), closer to our interest, present one aspect of the dynamics of sales. It is a competitive industry model, where consumers are assumed to choose a store based on the price of a single item, and firms are informed about other firms' prices and, hence, sales. Jeuland and Narasimhan (1985) present a related idea in the context of a monopolist.

Dudine, Hendel, and Lizzeri (2006) provide an analysis of the role of commitment in a monopoly market with storable goods. They only consider linear prices and show that, in contrast with the literature on the Coase conjecture that discusses

---

[3] Skipping every second purchase is the most tempting deviation: if a consumer does not want to skip once, they also do not want to skip more than once.

durable goods markets, removing commitment leads to higher prices, rather than lower prices, when goods are storable and demand can be anticipated.

Nava and Schiraldi (2012) explain sales, in the absence of consumer heterogeneity or a discrimination motive, based on collusion. Sales, and the induced storage, lower the incentives to deviate from collusion, and lower payoff during punishment periods.

Although it does not involve storage, Sobel (1991) also presents a model of sales (see also Conlisk, Gerstner, and Sobel 1984; Sobel 1984; and Pesendorfer 2002). The model involves a market with a durable good monopolist; at every date a mass of new consumers enter the market. Consumers have unit demands and two possible valuations for the good. Sobel (1991) characterizes the set of equilibria under the assumption that the monopolist cannot commit. An important feature of the analysis is that there can be price cycles.

Price cycles are also generated in the customer recognition literature (Villas-Boas 2004) where firms price nonanonymously, according to previous purchasing behavior.

There is an extensive literature on durable goods.[4] The distinction between the products we consider, storables and durables, is tricky. The durable good literature is largely based on unit demand, one-time purchase, and the incentives to postpone such a purchase. In contrast, the focus of this paper is on storage, which permits anticipating purchases. Naturally, multiple units of durables, cars or TVs, are often consumed. Buyers of most durables return to the market, and may do so in anticipation of their needs should prices grant it. We view the distinction between the assumptions in the durable good literature and our paper, as one capturing the frequency of purchase. For infrequently purchased products the one-time, single unit purchase seems like a reasonable simplification. Instead, for frequently purchased nonperishable products, storage and demand anticipation are important forces to model. There is also a recent literature that studies the effect on market efficiency of the timing of trade in the context of markets with adverse selection. For instance, Fuchs and Skrzypacz (2012) show that allowing frequent trading can worsen the adverse selection problem. This is in contrast with Hendel, Lizzeri, and Siniscalchi (2005), who show that frequent trading can generate perfect sorting in a durable goods market with adverse selection, where new goods are constantly being produced and then depreciate.[5]

## II. Two-Period Model

### A. *Setup*

We first consider a two-period model with homogenous consumers. The simplicity of the two-period environment helps to highlight some basic forces created by the interplay between storability and nonlinear pricing.

Buyers' per period willingness to pay for consumption $C$ is denoted $V(C)$. The function $V(C)$ is assumed to be increasing, concave, and continuously differentiable, with $V(0) = 0$ and a saturation point $C^*$, so that for all $C \geq C^* : V(C) = V(C^*)$

---

[4] See Waldman (2003) for a survey.
[5] These papers feature common value uncertainty and no producer with monopoly power. The focus of these papers are very different from that of the present paper.

and for all $C < C^* : V(C) < V(C^*)$. The latter guarantees bounded monopoly profits.[6] We normalize the marginal cost of production to zero.

As a benchmark, note that absent storability, the seller would offer the efficient quantity $C^*$ and extract all the surplus with tariff $P^* = V(C^*)$ each period.

For simplicity we assume that there is no discounting. Thus, if the consumer has $Q$ units to allocate over the two periods, he would like to consume equal amounts in each period. Consumers' payoffs from consumption $C_1$ and $C_2$ and payments $P_1$ and $P_2$ is given by

$$V(C_1) \; + \; V(C_2) \; - \; (P_1 \; + \; P_2)$$

We assume that all consumers have a storage capacity of $S$. Storing quantity $s$, so that $0 \leq s \leq S$, is free, but it is impossible to store more than $S$.[7] The feasible consumption of a consumer who purchases bundles $Q_1$ and $Q_2$ in period 1 and 2, respectively, is given by $C_1 = Q_1 - s$ and $C_2 = Q_2 + s$.

Finally, we assume that the monopolist can commit to future actions. In the first period, the monopolist announces (and commits to) the menu for each period, and then consumers make decisions.[8] All transactions take place on the spot market, so that past history of purchases has no effect on current transactions.[9] The storage capacity is known by the seller. We relax this assumption later, when we allow heterogeneous storage.

### B. *Optimal Policy*

The monopolist's problem is to choose the sequence of bundles $Q_1$ and $Q_2$, and tariffs $P_1$ and $P_2$ to maximize the sum of revenues $P_1 + P_2$. The transfers are determined by three constraints.

First, we impose the standard participation constraint modified to take into account the possibility of storage. Transfers must be lower than the sum of utilities:

(1) $$\max_{0 \leq s \leq S} \{V(Q_1 - s) - P_1 + V(Q_2 + s) - P_2\} \geq 0.$$

Second, consumers must be willing to purchase both bundles, as opposed to just one. Note that in the absence of storage, consumption in period 1 is independent of the consumption in period 2 and, hence, the participation constraints for period 1

---

[6] Alternatively, we could assume a positive marginal cost and $\lim_{C \to \infty} V'(C) = 0$.

[7] An earlier version of the paper considered a convex cost of storage and results were qualitatively similar.

[8] We briefly discuss the case of no commitment in the two-period model in the conclusions. Much of the durable good literature focuses on the case in which the seller lacks commitment, partly because the commitment solution is straightforward. Understanding lack of commitment in the context of our model would be interesting (albeit complex). However, commitment is a useful benchmark to which the case with no commitment must be compared to understand the role of commitment. Furthermore, this is not an implausible assumption for some of the markets we model, in part due to the frequency in which consumers return to the market (in the case of retailers) and the frequency with which manufacturers and retailers contract, say, in the case of supermarkets.

[9] If one allows the monopolist to bundle the portions of the good delivered in the future, the problem becomes trivial: the monopolist can offer the contract under which he delivers the efficient level of consumption in each period and the consumer pays his full surplus to the monopolist.

and 2 are independent. By allowing the consumer to store, we introduce an additional obstacle for second-period participation. The consumer's outside option in the second period is greater than zero, due to storage.

The constraint is that the consumer should be willing to purchase $\{Q_2, P_2\}$ in the second period, rather than purchase only $\{Q_1, P_1\}$ in the first period, and optimally smooth $Q_1$ over the two periods:

$$(2) \qquad \max_{0 \le s \le S} \{V(Q_1 - s) - P_1 + V(Q_2 + s) - P_2\}$$

$$\ge \max_{0 \le s \le S} \{V(Q_1 - s) - P_1 + V(s)\}.$$

Finally, the consumer should be willing to purchase $\{Q_1, P_1\}$ in the first period:

$$(3) \qquad \max_{0 \le s \le S} \{V(Q_1 - s) - P_1 + V(Q_2 + s) - P_2\} \ge V(Q_2) - P_2.$$

The next Lemma shows that only two of these three constraints are binding.

LEMMA 2.1: *Constraints* (2) *and* (3) *imply constraint* (1).

For all proofs we refer the reader to the Appendix.

We call constraints (3) and (2) no-skipping constraints. They are participation constraints that assure the consumer does not skip a purchase event in any period. The simplicity of the two-period setup stems from the limited forms of skipping. In a longer horizon the consumer can prepare herself to skip in numerous ways, namely, the consumer can store out any of the previous purchases as long as she has available storage. Thus, the longer the history, the more constraints on participation need to be imposed. As we show below, eliminating or limiting the ability to skip plays an important role in shaping optimal offerings by the monopolist.

We now characterize the optimal solution. Denote the optimal sales by $X_1$ and $X_2$ and recall that $C^*$ is the efficient consumption obtained by $V'(C^*) = 0$.

THEOREM 1: *Assume $S$ is such that $0 < S < C^*$.*[10] *In an optimal policy, the monopolist chooses first-period output $X_1 = C^* + S$, and second-period output $X_2$, such that $X_2 + S < C^*$. In this optimal policy:*

(i) *the monopolist induces a binding storage constraint for the consumer;*

(ii) *first period consumption is efficient;*

(iii) *second period consumption is below the efficient level; and*

(iv) *consumers enjoy positive surplus.*

---

[10] If $S \ge C^*$, the optimal policy for the monopolist is to set $X_1 = 2C^*$, $X_2 = 0$, and $T_1 = 2V(C^*)$: when storage capacity is very large, the monopolist can extract all the surplus by selling everything in the first period. We have in mind products that are frequently purchased, so that selling once and for all is not relevant nor interesting.

Storability induces the monopolist to create distortions in consumption away from the first best, even though consumers are homogenous and the monopolist can choose any nonlinear price. Furthermore, the monopolist optimally chooses to allow the consumer to enjoy positive surplus. The two new no-skipping constraints introduced by storability are tighter than the standard participation constraint: if the monopolist targets all the surplus, the consumer always has the option of not purchasing in the second period and smoothing consumption through storage. However, this is not the whole story, since the seller could generate an allocation with fully extractable surplus (for example, by selling only in the first period).

To make the skipping of period-two purchases more difficult, the seller could lower first period quantity.[11] Alternatively, the seller could increase first period quantity beyond the capacity constraint. Indeed the monopolist's optimal policy is designed in such a way, that the consumer fills his storage up to its maximum capacity. Clearly, the storage provides an additional freedom for the consumer to allocate the consumption more efficiently across time, and, hence, gives him bargaining leverage against the monopolist. However, once the storage is filled, any additional quantity sold in the first period must be consumed immediately, and, hence, the monopolist can extract the full value of these additional units. This is the underlying reason why the monopolist induces a binding storage constraint for the consumer. Moreover, this also explains why first-period consumption is efficient. Since the seller can extract all surplus (as no first-period quantity can be stored because storage is already filled) from these additional units, it is optimal to expand consumption up to the efficient level.

Thus, the seller offers $X_1 = C^* + S$ and is able to extract the following first-period tariff:

$$P_1 = \underbrace{V(X_1 - S)}_{C^*} + \underbrace{V(X_2 + S)}_{C_2} - V(X_2).$$

This amount equals the full extra surplus the consumer obtains from the bundle $X_1$, where $V(X_1 - S)$ is the surplus of consuming $X_1 - S$ in the first period and $V(X_2 + S) - V(X_2)$ is the surplus of consuming an extra $S$ on top of the bundle $X_2$ in the second period.

Let us now check the second-period offering and why the consumer is left with a positive surplus. The most the seller can extract in the second period is the difference in utility between showing up in both periods and buying only in the first

---

[11] If the storage capacity is small enough, the monopolist has an alternative optimal policy, in which the role of the periods is reversed but the same profit is attained. The second period is the one in which the consumption is efficient. Intuitively, the no-skipping constraint becomes less restrictive when the storage capacity is small enough, so that the monopolist can make sure that the consumer is willing to purchase a large bundle (and pay a high tariff) in the second period. In this case, it is still true that the first-period tariff extracts the full surplus that is generated by the first-period bundle. We have downplayed this second implementation for the following reasons. First, we are interested in the impact of storage, so we want to focus on the case where storage is large (bigger threat). Second, as we will see in the next section, the alternative equilibrium is not robust to the introduction of heterogeneity in storage.

period. Since storage is binding, the latter is $V(X_1 - S) + V(S)$, the efficient consumption plus full storage consumption. The second period tariff is

$$P_2 = \underbrace{V(X_1 - S)}_{C^*} + \underbrace{V(X_2 + S)}_{C_2} - \underbrace{V(X_1 - S) - V(S)}_{C^*} = \underbrace{V(X_2 + S) - V(S)}_{C_2}.$$

We are now ready to discuss the intuition behind the optimal $X_2$. Second-period consumption is not efficient for the following reason. The second-period tariff increases with the size of the bundle $X_2$, so it seems that the seller would have an incentive to sell enough in the second period to lead to a second-period consumption $X_2 + S = C^*$. Indeed the latter maximizes $P_2$. However, these additional second-period units affect the amount that can be extracted in the first period. Recall that the first period tariff consists of two parts: the value of consumption in the first period and the additional value of consuming the stored inventories in the second period, $V(X_2 + S) - V(X_2)$. Since the consumer's valuation function is concave, this part of the first-period tariff is decreasing in $X_2$, so that second-period consumption is a threat to first-period surplus extraction. The monopolist has to balance the two effects, and ends up selling a bundle that leads to less than efficient second period consumption.

Let us return to surplus extraction. Both tariffs are set so that the seller captures the extra surplus generated by each offering. In other words, as we saw above, $P_1$ captures the additional surplus generated by $X_1$ relative to the consumer's utility should she only enjoy $X_2$ in the second period. This is similar for $P_2$. The seller is able to capture all additional surplus from each bundle. The reason the consumer manages to keep a positive surplus is that the marginal impact of $X_1$ is evaluated at $X_2$ and vice versa. Since $V$ is concave, the marginal surplus of each offering is less than the surplus of removing both bundles.

### C. *Heterogeneous Storage Capacities*

We now extend the model by allowing for a limited amount of consumer heterogeneity. The standard second-degree price discrimination model (see, for instance, Tirole 1988, ch. 1) characterizes optimal nonlinear prices when there is heterogeneity in consumer valuations. One could of course perform the same exercise within our environment. However, we think that it is more useful to keep our focus on the effects of storability, so we retain the assumption of homogeneous consumer valuations, and instead allow consumers to have heterogeneous storage capacities. We assume that a fraction $\alpha$ of consumers has no ability to store, and a fraction $1 - \alpha$ has storage capacity $S$, as above.

We reserve capital letters for the consumer with storage $P_1, P_2, Q_1$, and $Q_2$ and small letters $p_1, p_2, q_1$ and $q_2$ for the consumer without one. For brevity we will call the consumers with storage S, and the consumers without storage NS.

The monopolist maximizes profits given by

$$\pi = \alpha(p_1 + p_2) + (1 - \alpha)(P_1 + P_2)$$

subject to the constraints that guarantee consumers choose the bundles that are meant for them. All the constraints that appeared in the case of single consumer carry over to this case, but we now must impose new self-selection constraints.

Neither type of consumer should skip a purchase in either period. For the storing consumer, it amounts to constraints (3) and (2) above. For the NS-consumer, they amount to the usual static participation constraints:

$$(4) \hspace{4cm} p_1 \leq V(q_1)$$

$$(5) \hspace{4cm} p_2 \leq V(q_2).$$

Furthermore, NS-consumers should not prefer to switch to the bundles $P_1, Q_1$, or $P_2, Q_2$:

$$(6) \hspace{3cm} p_1 - P_1 \leq V(q_1) - V(Q_1)$$

$$(7) \hspace{3cm} p_2 - P_2 \leq V(q_2) - V(Q_2);$$

and S-consumers should not switch to the whole bundle meant for NS-consumers:

$$(8) \quad p_1 + p_2 - P_1 - P_2 \geq \max_{0 \leq s \leq S} \{V(q_1 - s) + V(q_2 + s)\}$$
$$- \max_{0 \leq s \leq S} \{V(Q_1 - s) + V(Q_2 + s)\}.$$

S-consumers should not substitute any part of their bundle with the deal that is offered to NS consumers:

$$(9) \hspace{2cm} p_1 - P_1 \geq \max_{0 \leq s \leq S} \{V(q_1 - s) + V(Q_2 + s)\}$$
$$- \max_{0 \leq s \leq S} \{V(Q_1 - s) + V(Q_2 + s)\}$$

$$(10) \hspace{2cm} p_2 - P_2 \geq \max_{0 \leq s \leq S} \{V(Q_1 - s) + V(q_2 + s)\}$$
$$- \max_{0 \leq s \leq S} \{V(Q_1 - s) + V(Q_2 + s)\};$$

and, finally, S-consumers should not prefer to choose just one period of the bundle intended for NS-consumers:

$$(11) \hspace{2cm} p_1 - P_1 - P_2 \geq \max_{0 \leq s \leq S} \{V(q_1 - s) + V(s)\}$$
$$- \max_{0 \leq s \leq S} \{V(Q_1 - s) + V(Q_2 + s)\}$$

$$(12) \hspace{1cm} p_2 - P_1 - P_2 \geq V(q_2) - \max_{0 \leq s \leq S} \{V(Q_1 - s) + V(Q_2 + s)\}.$$

The next result offers a characterization of the optimal solution, in this case with two types of consumers.

THEOREM 2: *The optimal bundles with heterogeneous storage are such that* $q_1(\alpha) < C^*$, $q_2(\alpha) = C^*$, $Q_1(\alpha) = C^* + S$ *and* $Q_2(\alpha) < C^* - S$. *Under the optimal policy:*

   (i)  *the monopolist induces a binding storage constraint for the S-consumers;*

  (ii)  *S-consumers' consumption is efficient only in the first period;*

 (iii)  *NS-consumers' consumption is efficient only in the second period; and*

 (iv)  *S-consumers enjoy positive surplus and there is $\hat{\alpha}$ such NS-consumers enjoy positive surplus if $\alpha < \hat{\alpha}$.*

The S-consumer's bundles $Q_1$ and $Q_2$ are qualitatively similar to those characterized in Theorem 1. The only difference is that the second period quantity is affected by the presence and proportion of nonstorers. The presence of S-consumers, instead, drastically changes the way NS-consumers' bundles are priced. Both consumption and surplus are affected. First period consumption drops below the efficient level, and even NS-consumers enjoy a positive surplus.

When facing heterogeneous consumers, the monopolist has to worry not only about participation and no-skipping constraints, but also about incentive compatibility constraints. To understand the way incentive compatibility affects the offerings suppose that the bundles for the S-consumer are priced as in Theorem 1. The first period bundle $Q_1 = C^* + S$ is sold. We showed that the S-consumer pays for this bundle more than his first period surplus, since he purchases additional quantity $S$ for second period consumption. Do NS-consumers prefers to buy the bundle $Q_1$ instead of $q_1$? They do not. All consumers have the same valuation for the good, but the NS-consumer unable to store, have to consume the whole bundle in the first period. They are willing to pay at most $V(C^*)$, which is lower than what S-consumer pays, $P_1$. However, the reverse is possible. If $q_1$ is large enough, the S-consumer can be tempted to buy the cheaper bundle of the two, which is $q_1$. The latter might be cheaper as it is priced for current consumption, but can be stored (and smoothed) by S-consumers.

Now consider the second period. NS-consumers are offered bundle $q_2 = C^*$. The S-consumer does not value this bundle as much because he has $S$ already in storage (recall, that $C^*$ is the saturation point). On the other hand, if the price for the bundle $q_2$ is close to $V(q_2)$, NS-consumer might prefer the smaller bundle $Q_2$ that sells for less than $V(Q_2)$, since it is priced for a consumer who has supply in storage.

To summarize, the monopolist only has to make sure S-consumers do not switch to the NS-bundle in the first period, and NS-consumers do not switch to S-bundle in the second period. We can now turn to NS-consumer tariffs:

$$p_1 = V(q_1)$$

$$p_2 = P_2 + V(q_2) - V(Q_2).$$

It is easy to see that the optimal policy extracts the full surplus from the NS-consumer in the first period, but the level of consumption is below the efficient one. In the second period, the situation is reversed: consumption is at the efficient level, but the tariff is lower than $V(C^*)$.

The monopolist can use two instruments to make sure consumers purchase their intended bundles: the induced consumption levels and the tariffs. It may be surprising that when the monopolist wants to keep the S-consumer from switching to the NS-bundle, he operates mostly by distorting consumption, and when he wants to prevent the NS-consumer from switching, he operates mostly through lowering the tariff. In the first period, the monopolist sets $q_1$ low enough so that it is not attractive for the S-consumer. Once the size of the bundle is determined, the monopolist can set the highest price under which the bundle is still purchased, which is $V(q_1)$. The monopolist cannot achieve the same effect only by using tariffs because S-consumers' effective willingness to pay in the first period is higher than NS-consumers', this is because S-consumers can smooth consumption of the good across periods, whereas NS-consumers cannot. Naturally, the higher the proportion of NS-consumer the more costly the quantity distortion is, that is why $q_1(\alpha)$ increases in $\alpha$.

Instead, in the second period, the S-consumer receives the bundle $Q_2$ at price $P_2 = V(Q_2 + S) - V(S) < V(Q_2)$. If the NS-consumer purchases $Q_2$ instead of $q_2$, he can get a surplus of $V(Q_2) - P_2$. The monopolist must guarantee the NS-consumer the same surplus from purchasing the bundle $q_2$. The only way he can eliminate this surplus is to set $Q_2 = 0$, which in turn will result in $P_2 = 0$. In the optimal solution, the monopolist sets the level of NS-consumption to be efficient and extracts all the surplus up to $V(Q_2) - P_2$, by setting a low enough tariff. As the proportion of NS-consumers increases, the seller is willing to further distort $Q_2$ downward, explaining why $Q_2(\alpha)$ is a decreasing function.[12]

The picture that emerges from this analysis is that the skipping constraints can be relaxed by filling the storage capacity of the consumer. Moreover, consumption flows of nonstorers are distorted downward to alleviate storers' incentives constraints.

Finally, recall that in the case of homogeneous consumers if the storage capacity is small enough, there exists an alternative optimal pricing scheme for the monopolist, in which the role of the two periods is switched: the monopolist induces efficient consumption in the second period, and distorts consumption of the good in the first period. When consumers are heterogeneous in storage capacity, this alternative pricing scheme is no longer optimal. This policy involves a small bundle in the first period, so that the consumer does not have significant outside options in the second period. However, if the small bundle is sold in the first period for the S-consumer, it must be the case that the small bundle is sold for the NS-consumer as well, which is a significant loss in the revenue collected from NS-consumers.

---

[12] If the share of the S-consumers is small enough, it is optimal to set $Q_2 = 0$ to make sure that all the surplus of the NS-consumers is extracted. In particular, if $\alpha \geq V'(S)/V'(0)$, it is optimal to set $Q_2 = 0$. However, for any value of $\alpha$, both types of consumers consume positive quantities.

### III.  Nonlinear Prices: Infinite Horizon

The two-period model has the advantage that it is a fairly simple way to gain an initial understanding of the constraints imposed by storability. The model is also simple enough that it could be extended to richer forms of heterogeneity. However, the two-period model also has some shortcomings. For example, it is unclear if the alternation between different levels of consumption, and the fact that the monopolist induces a binding storage constraint, are artifacts of the particular setup or fundamental features of storable goods. To address these questions, we approach the problem from another angle with some advantages and some other limitations. We consider a very stark infinite horizon model, in which the monopolist repeatedly interacts with the same consumers, so that the problem has some degree of stationarity. Of course, storage can endogenously introduce nonstationarities but the primitives are unchanging over time. In this model, consumers can start and end a period with storage, thereby enabling inventories as a tool at the disposal of the consumer in every period. This is not feasible in a two-period world. However, despite its extreme simplicity in some respects, the model is much more difficult to analyze. Because of this, we are only able to obtain a partial characterization of the monopolist's optimal policies.

We assume that time is continuous. This assumption is mostly made for technical reasons because it helps us to avoid dealing with divisibility issues.

The monopolist chooses a production and tariff schedule $(Y_t, R_t)$. $Y_t$ is interpreted as the total amount of good produced up to time $t$, and $R_t$ is the cumulative tariff that the agent has to pay to purchase $Y_t$. Naturally, we assume that $Y_t$ and $R_t$ are increasing and right-continuous in $t$. The exact meaning of the schedule $(Y_t, R_t)$ should become clear when we introduce the consumer's and the monopolist's maximization problems.

The monopolist maximizes the average profit, i.e.,

$$\limsup_{\mathcal{T} \to \infty} \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} b_t \, dR_t,$$

where $b_t \in \{0, 1\}$ is a decision by the consumer of whether to buy at time $t$.[13] We assume that the monopolist's production costs are zero.

As in the two-period model, we assume that the monopolist can commit to the sequence of bundles.[14] Again, we believe that this is a useful benchmark. Also, we

---

[13] In principle, we allow the monopolist to offer a menu $\{(Y_{i,t}, R_{i,t})\}_{i \in I \subset \mathbb{R}}$. In this section, we formulate the consumer's and monopolist's problems using a singleton menu to simplify the exposition. In full generality, the monopolist's objective function, for example, would be

$$\limsup_{\mathcal{T} \to \infty} \frac{1}{\mathcal{T}} \int_I \int_0^{\mathcal{T}} b_{i,t} \, dR_{i,t} \, di.$$

Since, in our model, social welfare is bounded and the monopolist's profit is bounded from below by zero, the integrals for the monopolist's profit and the consumer's utility converge for any individually rational strategy profile.

[14] However, given our assumptions, and in contrast with the two-period model, we believe that the optimal policy under commitment would also be an equilibrium as a limit of a suitably modified model without commitment. Note that in contrast with the extreme models of durable goods monopoly, in our model there are recurring sales, which gives ample scope for sustaining noncompetitive profits in equilibrium when discount factors are high.

do not allow the monopolist to offer prices that are contingent on individuals' history of purchases, so all transactions happen on an anonymous spot market.

The consumption of an agent is denoted by $c_t$. The flow value of consumption is given by $V(c_t)$. The consumer chooses his consumption stream $c_t$ and shopping schedule $b_t$ to maximize the average flow of utility

$$\lim_{T\to\infty} \sup \frac{1}{T} \int_0^T [V(c_t)\, dt - b_t\, dP_t],$$

where $V(c)$ is assumed to satisfy the same assumptions as in the previous section.

Note, that we assume that there is no discounting. This assumption, together with concavity of $V$, implies that an agent endowed with stock $Q$ of the good over time interval of length $T$, would like to consume $Q/T$ at each moment in time. This drastically simplifies the consumer's problem, allowing us to offer a particularly simple exposition of some of the key effects.

As in the two-period model, the consumer has storage capacity of size $S$. At each moment in time the consumer's inventories $S_t$ must satisfy

$$S_t = S_0 + \int_0^t [b_\tau\, dY_\tau - c_\tau\, d\tau]$$

$$0 \le S_t \le S.$$

The first equation states that the change in the consumer's inventories is equal to the difference between the amount purchased and the amount consumed. The second inequality is the capacity constraint.

To summarize, the consumer's maximization problem is

$$\max_{c_t, b_t} \left\{ \lim_{T\to\infty} \sup \frac{1}{T} \int_0^T [V(c_t)\, dt - b_t\, dR_t] \right\}$$

$$\text{s.t.} \quad S_t = S_0 + \int_0^t [b_\tau\, dY_\tau - c_\tau\, d\tau]$$

$$0 \le S_t \le S.$$

The monopolist's problem is

$$\max_{R_t, Y_t} \left\{ \lim_{T\to\infty} \sup \frac{1}{T} \int_0^T b_t^*(R, Y)\, dR_t \right\},$$

where $b_t^*(R, Y)$ is a consumer's best response.

We reserve small letters for flow variables and capital letters for all other ones.

It is important to distinguish between points of continuity and discontinuity of $Y_t$. We interpret these as flow and stock sales, respectively. A flow sale $q_t = dY_t/(dt + 0)$ is of the same magnitude as instantaneous consumption, which means that it is of measure zero relative to any interval of periods of positive

consumption. This of course does not mean that flow sales cannot be stored. If the consumer buys the flow amount $q_t$ at time $t$ he should pay $p_t = dR_t/(dt - 0)$. The stock sale $Q_t = Y_t - \lim_{\tau \to t - 0} Y_\tau$ is of the same magnitude as the consumer's inventory, and therefore it cannot be gainfully used for instantaneous consumption. If the consumer buys a stock amount $Q_t$, he has to pay $P_t = R_t - \lim_{\tau \to t - 0} R_\tau$.

Observe that if the consumer has no storage capacity, i.e., $S = 0$, then the optimal policy for the monopolist is to induce the efficient level of consumption $c_t = c^*$ for all $t$, and extract the entire surplus by charging $p_t = V(c^*)$ for all $t$. This policy corresponds to $Y_t = c^* t$ and $R_t = V(c^*) t$.

However, we now show that when the consumer can store, a monopolist who offers only flow bundles completely loses the power to price nonlinearly.

## A. Storability Eliminates Nonlinear Pricing of Flow Sales

The easiest way to see that no extra surplus can be extracted from nonlinear pricing of flow bundles is to consider the case in which the monopolist offers a constant flow $q_t = q$ at flow-bundle price $p$. Given this policy by the monopolist, consumers' optimal consumption is given by $V'(c) = p/q$ in all periods, implying that consumers can fully unbundle the monopolist's flow-bundle price.

More formally, denote by $\lambda$ the fraction of time when a consumer purchases the good. Since the consumer can accumulate inventories, and the capacity constraint is not binding in the case of flow sales, perfect smoothing of consumption is feasible and optimal, so that $c_t = \lambda q$ in every period. The cost of this policy, on average, is $\lambda p$ per period. Thus, the consumer picks $\lambda$ as follows:

$$\lambda^l = \arg \max_{\lambda \in [0, 1]} \{V(\lambda q) - \lambda p\}.$$

If the per unit price is not too high (i.e., $p/q < V'(0)$, otherwise there would be no purchases), the optimal $\lambda^l$ satisfies

$$V'(\lambda^l q) = \frac{p}{q}.$$

Flow profits in this case are given by $c^l V'(c^l)$, where $c^l = \lambda^l q$. Notice that $c^l$, by the first-order condition above, is also the optimal consumption of a buyer that faces the linear price $p/q$. Thus, with storage, selling a flow bundle $q$ at bundle price $p$ is equivalent to setting a linear price $p/q$.

The implication is that when the monopolist offers a constant flow of bundles $q$, the ability of the consumer to store destroys all the monopolist's ability to price nonlinearly: the best policy within this class is equivalent to a static linear pricing policy. This result is an intertemporal parallel to the common wisdom that successful nonlinear pricing requires constraints on arbitrage among consumers. In our model, the arbitrage takes place across the different periods for the same consumer.

The next results show that the difficulty faced by the monopolist is much more pervasive. We now consider what happens for any sequence of flow-bundle sales, not just stationary ones.

THEOREM 3: *If* $S > 0$ *and the monopolist is restricted to sell only flows of the good* (*i.e.*, $Y_t$ *is continuous*), *then the optimal policy is revenue-equivalent to linear pricing.*

Although the proof is more complex, the logic of this result is similar to the one outlined above for constant flow-bundle sales: with flow sales, the storage constraint is never binding and the consumer can time his purchases to unbundle the monopolist's attempt to price bundles nonlinearly.

To develop the intuition behind Theorem 3, let us go through the main features of the optimal consumption and purchasing sequences. First, note that the optimal consumption stream is constant whenever the storage constraint is not binding. This follows directly from the concavity of $V$. Second, more importantly, the consumer buys the good from the monopolist at time $t$ only if

$$V'(c_t) \geq \frac{p_t}{q_t}.$$

If this condition does not hold, it is cheaper for the consumer to trade with his past self using a linear tariff, i.e., to relocate some of his consumption from a previous moment in time using storage, rather than to buy from the monopolist. This condition puts an upper bound on the price that can be charged by the monopolist. If we rewrite it as

$$q_t V'(c_t) \geq p_t,$$

and, roughly speaking, integrate both sides of this inequality, we will get an average profit of the monopolist on the right-hand side, and $c_t V'(c_t)$ on the left-hand side. The latter is equivalent to a profit from a linear tariff.

Finally, to obtain the full characterization of the optimal consumption stream, one needs to make sure that the boundary conditions for this problem are satisfied. In particular, it must be that the total amount of good purchased (following the rule discussed in the previous paragraph) equals to the total amount consumed.

## B. *Storability and Periodic Sales*

We now investigate whether the monopolist can restore some of its ability to extract surplus via nonlinear prices by resorting to stock bundles rather than just flow bundles. The next result shows that simply selling stocks instead of flow bundles is not sufficient: intertemporal arbitrage by the consumer does not depend on the monopolist selling a flow of goods, but rather depends on the frequent availability of purchasing opportunities.

Suppose the monopolist is restricted to offer a constant bundle $Q$ at (nonlinear) price $P$ at each point in time. Such a class of policies can be represented as menus $\{(Y_{i,t}, R_{i,t})\}_{i \in [0,1]}$ that satisfy

$$Y_{i,t} = \lfloor t + i \rfloor Q \quad \text{and} \quad R_{i,t} = \lfloor t + i \rfloor P.$$

The monopolist is free to choose $Q$ and $P$.

REMARK 1: *This class of policies is revenue-equivalent to linear pricing.*

Indeed, the consumer's problem is equivalent to choosing the intervals, $T$, at which to purchase in order to maximize flow utility. The consumer then optimally smooths consumption within the period, to get a utility flow $V(Q/T)$:

$$\max_T \left( V\left(\frac{Q}{T}\right) - \frac{P}{T} \right).$$

The optimal $T$ solves:

$$-\frac{1}{T^2}\left( QV'\left(\frac{Q}{T}\right) - P \right) = 0 \quad \text{or}$$

$$V'\left(\frac{Q}{T}\right) = \frac{P/T}{Q/T}.$$

The consumer times purchases so that marginal utility equals the unit price of flow consumption $(P/T)/(Q/T)$, where the numerator is flow price, and the denominator is flow purchases. Thus, with stationary policies, the monopolist loses all the ability to price nonlinearly. Profits are not higher than charging a linear tariff.[15]

The main force behind consumers' ability to intertemporally unbundle nonlinear prices is due to the fact that they have ample opportunities to time their purchases to construct their desired sequence of consumption.

We now show that, by choosing cyclical policies, the monopolist can do better than the profits from linear prices. It can partially restore its ability to extract surplus via nonlinear pricing by limiting the opportunities for consumers to time purchases and unbundle the nonlinear prices. The idea is to limit consumers smoothing opportunities by selling only infrequently, and forcing the consumer to buy bundles that use-up all the storage capacity, at each purchase.

We only consider the class of periodic sales where the monopolist offers a stock $Q$ at bundle price $P$ at periods separated by constant time intervals $T$; in all other periods the monopolist does not sell anything (or sets a price so high that consumers will never purchase).[16] Formally, the production and tariff schedule is defined as

$$Y_t = \left\lfloor \frac{t}{T} \right\rfloor TQ \quad \text{and} \quad R_t = \left\lfloor \frac{t}{T} \right\rfloor TP.$$

---

[15] We neglected the constraint $Q \leq S$. Selling more than $S$ would be immaterial, since purchases cannot be carried forward.

[16] We conjecture that the policy outlined in Theorem 4 is optimal in the class of all policies, but we could not prove this. This result does show that any optimal policy has to have cyclical characteristics because our periodic sales policy does better than any constant stationary policy.

THEOREM 4: *The monopolist can improve on linear pricing profits by using periodic sales.*

*In the class of periodic sales, the optimal policy is to only make available a bundle equal to the storage capacity S at periods separated by constant intervals $T^S$. The optimal price $P^S$ and interval $T^S$ are given by*

$$P^S = 2T^S\left(V(c^S) - V\left(\frac{c^S}{2}\right)\right); \qquad T^S = \frac{S}{c^S},$$

*where $c^S$ is the consumption that solves*

$$c^S = \arg\max_{c \geq 0}\left\{V(c) - V\left(\frac{c}{2}\right)\right\}.$$

*At this allocation, consumption, profits, and welfare are independent of S. In addition, if the function $cV'(c)$ is single-peaked, consumption, profits, and welfare are strictly between those obtained under linear pricing and those obtained under nonlinear prices absent storage.*

PROOF:
Without loss of generality we can assume that all bundles offered by the monopolist are actually purchased by the consumer at the optimum. Otherwise the monopolist can, at no loss, redesign the policy to get rid of the bundles that are not purchased.

Suppose that the monopolist sells bundle $Q$ every $T$ periods and charges $P$. Consumption in this case is $c = Q/T$. Since the consumer makes a purchase every $T$ periods, the price must be such, that he is not willing to skip a purchase and consume his inventories. The following inequality guarantees that the consumer does not wish to skip a single purchase:

(13) $$\frac{P}{T} \leq 2V(c) - 2V\left(\frac{c}{2}\right).$$

We now show that the above inequality implies that skipping more than one purchase in a row is not beneficial either. To show that, we need to prove that

$$k\frac{P}{T} \leq (k+1)V(c) - (k+1)V\left(\frac{c}{k+1}\right).$$

From (13), we know that

$$k\frac{P}{T} \leq 2kV(c) - 2kV\left(\frac{c}{2}\right).$$

Note, that

$$2kV(c) - 2kV\left(\frac{c}{2}\right) - \left((k + 1)V(c) - (k + 1)V\left(\frac{c}{k + 1}\right)\right)$$

$$= (k - 1)V(c) + (k + 1)V\left(\frac{c}{k + 1}\right) - 2kV\left(\frac{c}{2}\right).$$

By concavity of $V(\cdot)$, we obtain that

$$(k - 1)V(c) + (k + 1)V\left(\frac{c}{k + 1}\right) \le 2kV\left(\frac{c}{2}\right),$$

hence,

$$k\frac{P}{T} \le 2kV(c) - 2kV\left(\frac{c}{2}\right) \le (k + 1)V(c) - (k + 1)V\left(\frac{c}{k + 1}\right).$$

Of all constraints for this problem, (13) is the tightest. Notice, that (13) does not depend on $Q$, but only depends on consumption. Hence, setting $Q^S = S$ is weakly optimal for the relaxed problem, which only involves the constraint that we discussed above. Also, by setting $Q^S = S$, we make all other constraints obsolete. In fact, if we set $Q^S = S$, consumers cannot use inventories that were purchased more than $T$ periods ago for current consumption. Hence, we can restrict our attention to events of skipping consecutive purchases.

Since the monopolist is maximizing the flow of payments, he can set $P/T = 2V(c) - 2V(c/2)$ and solve for the optimally induced consumption:

$$c^S = \arg\max_{c \ge 0}\left\{V(c) - V\left(\frac{c}{2}\right)\right\}.$$

The rest of the solution is straightforward: $T^S = S/c^S$ and $P^S = 2T^S V(c^S) - 2T^S V(c^S/2)$.

It remains to show that both profits and the induced consumption at the optimum are strictly between those that arise under linear pricing and in the absence of storage. If the monopolist sells a bundle $S$ at constant intervals $T$ and induces the consumption $c$, the flow profit is $\pi^S = 2V(c) - 2V(c/2)$. If storage is not feasible, the optimal consumption satisfies $V'(c^*) = 0$ and the flow profit is $\pi^* = V(c^*)$. If the monopolist uses the optimal linear pricing that induces the consumption $c^l$, the flow profit is $\pi^l = c^l V(c^l)$.

First we prove that when storage is unavailable both profits and the consumption are larger then under optimal periodic policy. Note first that $2V(c) - 2V(c/2) < V(c)$ by the strict concavity of $V(c)$. Since $c^*$ maximizes $V(c)$, we obtain that $V(c^*) \ge V(c) > 2V(c) - 2V(c/2)$ for any $c$ including $c^S$. The first-order condition for $c^S$ is $V'(c^S) - V'(c^S/2)/2 > 0 = V'(c^*)$, hence, by strict concavity of $V(c)$, we obtain that $c^* > c^S$.

Under linear pricing, the largest profit the monopolist can obtain is $c^l V'(c^l)$. Here, we argue that this profit is achievable under the periodic policy as well. Indeed, observe that

$$2V(c^l) - 2V\left(\frac{c^l}{2}\right) = 2 \int_{\frac{c^l}{2}}^{c^l} V'(x) \, dx.$$

Since $V'(c^l) < V'(c)$ for any $c < c^{l,}$ we get

$$2 \int_{\frac{c^l}{2}}^{c^l} V'(x) \, dx > 2 \int_{\frac{c^l}{2}}^{c^l} V'(c^l) \, dx = c^l V'(c^l).$$

Finally, $2V(c^S) - 2V(c^S/2) \geq 2V(c) - 2V(c/2)$ for any $c$, hence, $2V(c^S) - 2V(c^S/2) > c^l V'(c^l)$.

To prove that $c^S > c^l$, we use the assumption that $cV'(c)$ is single-peaked. Observe, that the first-order condition for $c^S$ is

$$c^S V'(c^S) = \frac{c^S}{2} V'\left(\frac{c^S}{2}\right).$$

By single-peakedness of $cV'(c)$, it must be the case that $c^l \in (c^S/2, c^S)$. This completes the proof.

The policy outlined in Theorem 4 is cyclical. This gives us a theory of sales, based on storability, even in an environment with identical consumers and no time varying preferences or new entry of consumers.[17] The intuition for the fact that a cyclical policy can restore some of the ability by the monopolist to extract surplus via nonlinear prices is the following. The fact that the bundles are only available infrequently implies that there are fewer no-skipping constraints for the monopolist to worry about; it is more costly for the consumer to skip a purchase. By selling bundles that fill up the storage capacity of the consumer, the monopolist makes it harder for the consumer to smooth consumption in the event that he chooses to skip a purchase, thereby enabling the monopolist to extract more surplus.

To gain some additional insight into the role of infrequent sales, let us go back to the no-skipping constraint, which determines prices. Any purchase has to give the consumer higher utility than skipping it and smoothing optimally out of storage. The utility from buying for two periods in a row, $2TV(Q/T) - 2P$, has to exceed the utility from smoothing the first purchase over both periods, $2TV(Q/(2T)) - P$. Using both terms, we compute the highest $P$ that guarantees participation:

$$P = 2T\left[V\left(\frac{Q}{T}\right) - V\left(\frac{Q}{2T}\right)\right].$$

---

[17] Of course, one limitation is that we are not able to characterize the fully optimal policy. However, the optimal policy cannot involve stationary policies.

TABLE 1—LINEAR VERSUS NONLINEAR PRICING

| Regime | Consumption | Profits | C. Surplus |
|---|---|---|---|
| Nonlinear no storability | 1 | 1/2 | 0 |
| Nonlinear with storability | 2/3 | 1/3 | 1/9 |
| Linear | 1/2 | 1/4 | 1/8 |

Skipping saves $P$ but causes the loss in utility derived from consuming $Q/2T$ instead of $Q/T$ over the $2T$ interval. Why is it optimal to fill the storage capacity (namely, setting $Q = S$)? By properly adjusting $T$, the single-skipping constraint just considered leads to identical profits for any $Q$ as long as $Q/T$ remains unchanged. The advantage from setting $Q = S$ comes from the fact that, when $Q < S$, there is additional storage capacity available to plan ahead of skipping a purchase, thereby smoothing out more evenly the pain of skipping and inducing a tighter no-skipping constraint.

Recall that the first-order condition for consumption that maximizes flow revenues is $V'(c^S) = V'(c^S/2)/2$. Note first that $c^S < c^*$. This follows from $V'' < 0$ and the first-order conditions for $c^S$. If $V'(c^S) = 0$ and $V'(c^S/2) > 0$, the first-order conditions would not hold. The idea is that by increasing consumption toward the optimal level the seller generates more consumer surplus, but also increases the skipping threat. The higher $c^S$, the higher the utility from skipping the second purchase. This last effect pushes the optimal consumption below the efficient level. The last term is due to storage.

In order to compare $c^S$ to consumption under the optimal linear prices $c^l$ notice that the latter is set where revenue $cV'(c)$ is highest, namely, by $V'(c) + cV''(c) = 0$ or where the inverse demand elasticity $cV''(c)/V'(c) = 1$. In the optimal policy with periodic sales, instead, $2V'(c^S) = V'(c^S/2)$, which implies that the arc elasticity of $V'$ between $c$ and $c/2$ is 1. Under the standard assumption of decreasing elasticity (which is guaranteed by demand not being too convex), this implies that the elasticity evaluated at $c^S$ is larger than 1, and, thus, $c^S > c^l$, since the elasticity is 1 at the latter.

*Quadratic Example.*—In order to give some sense of the magnitude of the effects of storability, we now provide an example with quadratic preferences (linear demand). Assume $V(q) = q - q^2/2$. In this case, the optimal solution presented in Theorem 4 involves a frequency of sales, $T^* = 3S/2$. This is associated with a flow consumption $S/T^* = 2/3$ and average profits $P^*/T^* = 1/3$.

Table 1 offers a contrast between this solution and those of nonlinear pricing absent storability, and of linear pricing. As one can see, storability generates sizable distortions in consumption and a reduction profits relative to static nonlinear pricing, but periodic sales allow the monopolist to extract substantially more than via linear pricing.

## C. *Heterogeneous Storage Capacities*

As in Section IIC, we now consider the possibility that consumers are heterogeneous in their storage capacity. The purpose of this extension is to generate richer testable implications and more realistic pricing patterns. Indeed, an unpalatable feature in the policy outlined in Theorem 4 is that in between stock sales the monopolist does not sell.

We again assume that a fraction $(1 - \alpha)$ of consumers have storage capacity $S$ while the rest cannot store at all. All consumers have the same preferences.

The presence of no storage consumers (NS-consumers) reintroduces the necessity to offer flow bundles $q_t$. If this was the only option for the monopolist, then it would only capture the surplus from nonlinear pricing the nonstorers at the expense of reducing the surplus extracted from storers.

However, the monopolist can simultaneously offer a flow bundle intended for nonstorers and an infrequent stock bundle intended for storers. We now characterize the optimal policy under the assumption that the monopolist offers a flow bundle $q_t$ at price $p_t$ and a stock-bundle $S$ every $T$ periods at price $P$. This menu can be represented as $\{(Y_{s,t}, R_{s,t}), (Y_{f,t}, R_{f,t})\}$, where

$$ (14) \qquad\qquad Y_{f,t} = \int_0^t q_\tau \, d\tau $$

$$ (15) \qquad\qquad R_{f,t} = \int_0^t p_\tau \, d\tau $$

$$ (16) \qquad\qquad Y_{s,t} = \left\lfloor \frac{t}{T} \right\rfloor TQ $$

$$ (17) \qquad\qquad R_{s,t} = \left\lfloor \frac{t}{T} \right\rfloor TP. $$

NS-consumers only purchase flow $q_t$, and never stock $Q$, because the latter does not increase their consumption. S-consumers potentially can purchase both flow and bulk sales. Moreover, since the consumers are anonymous, the S-consumers can purchase multiple flow bundles at a time. So, if by the time $t$, the S-consumer decides to buy a measure $B_t$ of flow bundles, he will have to pay $\int_0^t [p_\tau dB_\tau + b_\tau dR_{s,\tau}]$, and his inventory will be

$$ S_t = S_0 + \int_0^t [q_\tau \, dB_\tau + b_\tau \, dY_{s,\tau} - c_\tau \, d\tau], $$

where $b_t$ is a binary decision whether to buy the stock bundle at time $t$ or not.

We first note that the price for the flow should be $p_t = V(q_t)$. If it were greater than $V(q_t)$, then consumers who cannot store would not buy the good. If it were less than $V(q_t)$, it could be raised up to $V(q_t)$ without violating the participation constraints of those who do not store, and at the same time make skipping bulk sales more difficult for those who have storage.

Recall, that if the monopolist offers a flow bundle $q_t$ to the NS-consumers alone, the flow profit that he collects is

$$ \pi^{NS}(q_t) = V(q_t). $$

As we showed in Section IIIB, the policy of selling a stock bundle of size $S$ to S-consumers every $T$ periods generates a flow profit of

$$\pi^S(c) = 2V(c) - 2V\left(\frac{c}{2}\right).$$

We also define an auxiliary function that is helpful in stating our next result:

$$\pi^\Delta(c, q) \equiv 2\left(V(x) - V\left(\frac{c}{2}\right) - \left(x - \frac{c}{2}\right)V'(x)\right)\mathbf{1}\left\{\frac{c}{2} \leq x\right\},$$

where

$$x(q) = V'^{-1}\left(\frac{V(q)}{q}\right).$$

We are now ready to state a result that characterizes the monopolist's optimal policy. As we mentioned, the monopolist can offer both stock bundles and flow bundles on the market. The only restriction we place on the monopolist is that if the stock bundles are offered, they have to be sold every $T$ periods, and the size of these stock bundles has to be constant. As we already mentioned, such policies can be represented as menus of the following form: $\{(Y_{s,t}, R_{s,t}), (Y_{f,t}, R_{f,t})\}$, where $Y_{f,t}$, $R_{f,t}$, $Y_{s,t}$, and $R_{s,t}$ are defined in (14), (15), (16), and (17). The monopolist is of course free to choose both the period length $T$ and the size of the bundle $Q$. He is also free to set any flow bundle sequence $(q_t, p_t)$.

THEOREM 5: *In the class of periodic policies, the optimal policy consists of both flow and stock bundles. A stock bundle of size $S$ is offered every $T^\alpha$ periods, and a flow bundle of constant size $q^\alpha$ is offered all the time. This policy induces a consumption $c^\alpha$ for the S-consumers and $q^\alpha$ for the NS-consumers. The pair $(c^\alpha, q^\alpha)$ solves*

$$(c^\alpha, q^\alpha) = \arg\max_{c,q}\left\{\alpha\pi^{NS}(q) + (1 - \alpha)\left(\pi^S(c) - \pi^\Delta(c, q)\right)\right\}.$$

*Under the optimal policy:*

(i) *S-consumers only purchase stock bundles of size $S$, and NS-consumers only purchase flow bundles of size $q^\alpha$;*

(ii) *S-consumers pay lower per-unit price than NS-consumers;*

(iii) *the presence of S-consumers lowers the consumption of NS-consumers: $q^\alpha \leq c^*$; and*

(iv) *the presence of NS-consumers increases the consumption of S-consumers: $c^\alpha \geq c^S$.*

The monopolist's profit consists of two parts. The first part is $\alpha\pi^{NS}(q) +$ $(1 - \alpha)\pi^S(c)$, which is the hypothetical profit the monopolist would collect if he

could perfectly identify which consumers can store and which cannot. The second part is $(1 - \alpha)\pi^\Delta(c, q)$. This is the penalty, due to the fact that the monopolist cannot prevent S-consumers from purchasing the flow bundles that are meant for NS-consumers. Indeed, if $V'(c/2) \geq V(q)/q = V'(x)$, the S-consumer's bargaining position improves since, in the event of skipping a sale of a stock bundle, he can purchase a flow bundle and increase his consumption by $(x - c/2)$. For the storage consumer, this increase in consumption is effectively priced according to a linear per unit price of $V'(x)$, which is smaller than the average per unit gain in utility $(V(x) - V(c/2))/(x - c/2)$. The gain from purchasing $(x - c/2)$ units at linear price $V'(x)$ cannot be extracted by the monopolist, hence must be granted to S-consumers as a discount on the price of a bulk bundle.

According to Theorem 5, S-consumers pay a lower per-unit price than NS-consumers. This result follows from the no-arbitrage condition. Since S-consumers have more freedom in the market (i.e., they can easily mimic the choices of NS-consumers), they pay a lower price.

Note that if $V'(c^S/2) < V(c^*)/c^*$, the optimal policy is a combination of two optimal policies for the S- and NS-consumers, as if the types of the consumers were observable. The condition above guarantees, that the flow bundles meant for NS-consumers are not going to be purchased by S-consumers, so the separation of types in this case comes for free.

In light of Theorem 3, it is worth emphasizing that the monopolist's policy involves both flows and stocks. Theorem 3 shows that flow sales can be linearized by the storing consumer, thus failing to deliver profits beyond linear pricing. It is interesting that, when consumers are heterogeneous in their storage capacities, Theorem 5 says that the monopolist benefits from offering stocks to storers, even when a flow is available. The monopolist has to leave as much surplus to storers as they would achieve by unbundling the flow offered to nonstorers. Actually, the flow is likely to be offered at a low linear price. Recall that the linearized price is $V(q)/q$. Moreover, since the ideal nonstorer bundle involves $V'(q) = 0$, the linearized price $V(q)/q$ is quite low, or at least the monopolist would like to target nonstorers with a bundle that leads to low linearized prices. The intuition for why it is still possible for the seller to extract more from storers than by just pricing linearly can be seen in a static framework. Can the seller benefit from selling a bundle to a consumer who can also purchase at linear price $p$? The seller can offer the efficient bundle, and price it to leave buyers with at least the surplus obtained from linear prices. A similar calculation is at work here, with the additional consideration that the seller offers less than optimal consumption to the storer due to the skipping constraint.

## IV. Extensions

### A. *Convex Production Costs*

We have assumed so far that the cost of production is linear. This assumption is less stringent, that may appear at first because even if the cost of production is convex, if the monopolist could freely store the good, then we would restore linearity of the effective cost of delivering the good.

However, if the monopolist has limited ability to store the good, then strict convexity of the cost of production would lead to some changes in the nature of the optimal solution. Unfortunately, studying a general case of convex costs of production is quite complex. Here, we present a very simple and particularly stark version of convexity. Assume that, at any date, the monopolist cannot produce more than $Q$ but produces freely up to $Q$.

If $Q < S$, then the no-skipping condition must be modified. Assume first that $S/2 < Q < S$. Given that there is slack in storage capacity, if the consumer skips, the consumer is better off smoothing consumption by gradually accumulating storage over a number of periods before skipping, so that he has full storage upon skipping. If the consumer stores an equal amount over $n$ periods, preceding the skipping period his payoff is

$$2TV\left(\frac{S}{2T}\right) - P - \lim_{n\to\infty} n\left(TV\left(\frac{Q}{T}\right) - TV\left(\frac{Q - \frac{S-Q}{n}}{T}\right)\right).$$

Taking limits as $n$ goes to infinity, we obtain

$$2TV\left(\frac{S}{2T}\right) - P - V'\left(\frac{Q}{T}\right)(S - Q).$$

Thus, the no-skipping condition is given by

$$2TV\left(\frac{Q}{T}\right) - 2P = 2TV\left(\frac{S}{2T}\right) - P - V'\left(\frac{Q}{T}\right)(S - Q),$$

and the optimal price is given by

$$P = 2T\left(V\left(\frac{Q}{T}\right) - V\left(\frac{S}{2T}\right)\right) + V'\left(\frac{Q}{T}\right)(S - Q).$$

If instead $Q < S/2$, then similar steps lead to

$$P = 2T\left(V\left(\frac{Q}{T}\right) - V\left(\frac{2Q}{2T}\right)\right) + V'\left(\frac{Q}{T}\right)(2Q - Q) = V'\left(\frac{Q}{T}\right)Q,$$

showing that in this case profits are the same as under linear pricing.

It is easy to show that profits are increasing as $Q$ goes from $S/2$ to $S$.

## B. *Storage Depreciation*

So far we have made the extreme assumption that storage does not depreciate.

It is natural to consider some form of storage depreciation and investigate how that changes our analysis. Depreciation should make the good closer to perishable. This suggests that the monopolist should be able to extract more surplus when storage

depreciates. To give a particularly stark example, assume an extreme form of depreciation, such as in the case of dairies, where the good gives the same value to consumers until some expiration date $e$, and then it becomes worthless. If the monopolist can easily store the good, then one possible option for the monopolist is to only sell the good when it is very close to the expiration date, making it effectively nonstorable. We will now assume that this option is not feasible, or that it is very costly for the monopolist to deliver goods with very short expiration dates. From now on, the expiration date is the effective one after the monopolist has made the good available to consumers.

Consider first the case where the expiration duration $e$ is low. Specifically, consider the socially efficient consumption $c^*$, and let $\tau$ be such that $\frac{S}{\tau} = c^*$. If $\tau > e$, then the monopolist can restore the optimal solution when the good is not storable: full efficiency and surplus extraction. To see this, let $Q = ec^*$. Let the monopolist sell $Q$ every $e$ periods and charge $P = eV(c^*)$. If the consumer skips a purchase, his utility over the skipped interval is zero because the product expires after the skipping date given the policy that we have specified. Thus, in this case, storage no longer undermines the monopolist's ability to extract surplus. Note that here the storage capacity is not binding, so it plays no role in the solution as long as $\tau > e$.

Assume now that $\tau < e$, so that expiration duration is longer. It is now no longer possible to extract the full surplus. Indeed, when $e$ is sufficiently large, we get back to the solution with no expiration; all that is required is that $e > 2T^S$, i.e., the length of the skipping interval under the optimal monopoly solution with no expiration.

Naturally, this example shows that with depreciation, storage capacity $S$ affects the monopolist solution.

## V. Concluding Remarks

We studied the impact of product storability on nonlinear pricing. We showed that storability can enable consumers to undo sellers' attempts to price nonlinearly. The constraint is particularly severe under constant offerings, in which case sellers cannot extract more surplus than under linear prices.

Cyclical pricing, in the form of infrequent and bulky sales, constrain buyers ability to undo nonlinear prices. Infrequent sales limit skipping opportunities, while bulky sales make it harder for consumers to get ready to skip a purchase. Thus, the model delivers a theory of sales. Unlike most explanation of sales, this one is **not** based on a discrimination motive. Allowing for heterogeneity in storage delivers testable implications consistent with observed patterns. Specifically, the policy outlined in Theorem 5 has some interesting empirical content. It predicts that sales are more relevant for large bundles. As discussed at the start of this paper, this is a typical pricing pattern found in scanner data. The result is also consistent with smaller, more urban-located stores to have less prominent promotional activity. Buyers in urban areas have less storage, or at least, a smaller proportion of them are likely to store. Naturally, we have to be careful making a strong empirical prediction due to the simplicity of the model, which allows for a very limited form of heterogeneity. It would be natural to assume that customers differ not only in storage but preferences as well, perhaps in a systematic way, with more intense buyers being more

likely to store. The positive correlation between storage and usage may reinforce the finding that larger containers are more likely to be promoted. However, this paper has only begun to explore the interactions of heterogeneity and storability. A richer model would be particularly useful for delivering a broader set of empirical predictions. The richer model could allow heterogeneous preference, correlation between preferences and storage, as well as other storage technologies.

There is a connection between our results and results in the dynamic agency literature on the limits on contracting imposed by the agent's ability to time effort or savings decisions to undermine complicated nonlinear incentive schemes. For instance, see Cole and Kocherlakota (2001), where savings impose significant constraints and Holmstrom and Milgrom (1987), where, under exponential utility and Brownian motion, and utility only over final consumption, the optimal contract is linear in aggregate final outcomes.

This paper has also focused on an environment where the monopolist that can commit to a sequence of bundles. One natural next step would be to consider the consequences of lack of seller commitment. In the two-period model, for instance, it is easy to see that the optimal policy characterized in Theorem 1 is **not** an equilibrium absent commitment. In any pure strategy equilibrium absent commitment, second period consumption has to be efficient, otherwise the monopolist could increase profits in the second period. It turns out that equilibrium without commitment is quite complicated to characterize even in the two-period problem. For instance, there is no pure strategy equilibrium in the two-period model without commitment: for a wide range of first-period output levels, it cannot be the case that in the second-period storage is known and identical for all consumers. This creates many complications but also raises some interesting question for possible follow-up work.

## Mathematical Appendix

PROOF OF THEOREM 1:

Observe that constraints (3) and (2) can be written as

$$P_1 \leq \max_{0 \leq s \leq S} \{V(Q_1 - s) + V(Q_2 + s)\} - V(Q_2)$$

$$P_2 \leq \max_{0 \leq s \leq S} \{V(Q_1 - s) + V(Q_2 + s)\} - \max_{0 \leq s \leq S} \{V(Q_1 - s) + V(s)\}.$$

Since the monopolist maximizes $P_1 + P_2$, both constraints are going to be binding at the optimum. Substituting the constraints into the objective function, the monopolist's problem becomes

$$\max_{Q_1, Q_2} \left\{ 2 \max_{0 \leq s \leq S} \{V(Q_1 - s) + V(Q_2 + s)\} - V(Q_2) - \max_{0 \leq s \leq S} \{V(Q_1 - s) + V(s)\} \right\}.$$

Suppose that the monopolist induces a binding storage constraint for the consumer. Lemma A.1 on page 28 proves that under the optimal policy this is indeed the case.

Note that, in our candidate policy, the consumer purchases in both periods. Therefore, the monopolist maximizes

$$\prod^{S} = \max_{Q_1, Q_2} \{V(Q_1 - S) + 2V(Q_2 + S) - V(Q_2) - V(S)\}.$$

At the optimum we must have

(A1) $$V'(Q_1 - S) = 0$$

and

(A2) $$2V'(Q_2 + S) \leq V'(Q_2).$$

Equation (A1) implies that the optimal first-period output is $X_1 = C^* + S$.

Let us now consider the optimal choice of second-period output. If $2V'(S) < V'(0)$, then the solution is $X_2 = 0$. Thus, the optimal solution involves $X_2 = 0$ whenever $S \geq \hat{S}$, where $\hat{S}$ solves

$$2V'(\hat{S}) = V'(0).$$

In this case it is immediate that second-period consumption is equal to $S < C^*$. From now on we assume that $S < \hat{S}$, so that $X_2$ must satisfy (A2) as an equality. We now show that $X_2 < C^*$ for every $S$, despite the fact that $Q_2 \geq C^*$ would satisfy the necessary condition given by equation (A2) (since $V'(Q) = 0$ for all $Q \geq C^*$). To see this, assume by way of contradiction that $X_2 \geq C^*$. Note that when $Q_2 \geq C^*$,

$$2V(Q_2 + S) - V(Q_2) = V(C^*) \equiv V^*.$$

However, for any $S > 0$,

$$\max_{Q_2} \{2V(Q_2 + S) - V(Q_2)\} \geq 2V^* - V(C^* - S) > V^*,$$

which is the desired contradiction.

Because $X_2 < C^*$, the optimality condition (for $S < \hat{S}$) is given by

$$2V'(X_2 + S) = V'(X_2) > 0.$$

This implies that $X_2 + S < C^*$, which concludes the proof of parts (ii) and (iii). For future reference, profits are

$$\prod^{S} = V^* + 2V(Q_2 + S) - V(S) - V(Q_2).$$

Finally the following lemma concludes the proof of this theorem.

LEMMA A.1: *Under the optimal policy, the storage constraints are binding.*

PROOF:
See supplementary Appendix, available at http://www.ucl.ac.uk/~uctpnr1/Nonlinear_app.pdf.

PROOF OF THEOREM 5:
We begin this proof with the assumption that the monopolist sets a constant size of a flow bundle as a part of the optimal policy. Once we find the characterization of the optimal policy under this assumption, we prove that, indeed, the monopolist can not increase the profit by allowing the flow bundle to change over time (see Lemma A.2 on page 31).

Let us look at the problem of the consumer with storage. Suppose, in addition to the stock bundle of size $S$, S-consumers purchase the flow bundle $\lambda$ share of time. The share of time $\lambda_1$ that they buy the flow bundle if they do not skip any bulk sales is defined by

$$V'(c + \lambda_1 q) \leq \frac{V(q)}{q},$$

and the share of time $\lambda_2$ that they buy the flow bundle if they skipped one bulk sale is defined by

$$V'\left(\frac{c}{2} + \lambda_2 q\right) \leq \frac{V(q)}{q}.$$

These conditions hold as strict inequalities if corresponding $\lambda$ is zero, i.e. if the per unit price for the flow is so high that we have a corner solution.

If both $\lambda_1$ and $\lambda_2$ are strictly positive, given the optimal menu, there should be no bulk sales. To show this, observe that, if both $\lambda_1$ and $\lambda_2$ are strictly positive, then the inequalities above, should hold as equalities, i.e.,

$$V'(c + \lambda_1 q) = \frac{V(q)}{q}$$

$$V'\left(\frac{c}{2} + \lambda_2 q\right) = \frac{V(q)}{q}$$

and, hence $c + \lambda_1 q = c/2 + \lambda_2 q$. Denote by $x(q)$ the consumption that is induced by linear pricing:

$$x(q) = V'^{-1}\left(\frac{V(q)}{q}\right).$$

The no-skipping constraint in this case becomes

$$\frac{P}{T} \leq 2\left(V(c + \lambda_1 q) - V\left(\frac{c}{2} + \lambda_2 q\right) + (\lambda_1 - \lambda_2)V(q)\right) = -\frac{c}{2q}V(q) \leq 0,$$

hence, setting $P = 0$ and $c = 0$ is constrained optimal. The monopolist's profit under the policy without the stock bundles is

$$\pi^q(q) = \alpha V(q) + (1 - \alpha)x(q)V'(x(q)).$$

Now we proceed to the case where bulk sales are present in the optimal menu. We first observe that $\lambda_1 \leq \lambda_2$, hence, it must be in this case that $\lambda_1 = 0$.

We have two cases to consider:

(i) $\lambda_2 = 0$ and

(ii) $\lambda_2 \in (0, 1)$.

Suppose that $\lambda_2 \in (0, 1)$. Then,

$$V'\!\left(\frac{c}{2} + \lambda_2 q\right) = \frac{V(q)}{q}.$$

Note that $c/2 + \lambda_2 q = x(q)$.

Given this, the monopolist's profit becomes

$$2(1 - \alpha)V(c) + \alpha V(q) - 2(1 - \alpha)\,(V(x(q)) - \lambda_2 V(q))\,.$$

If we substitute $\lambda_2 q = x(q) - c/2$, then the monopolist's profit becomes

$$\alpha V(q) + (1 - \alpha)x(q)V'(x(q)) + 2(1 - \alpha)\left(V(c) - V(x(q)) - \frac{V'(x(q))}{2}(c - x(q))\right).$$

When deriving this expression, we assumed that $\lambda_2 > 0$. This assumption is valid only if the following inequality is satisfied:

(A3) $$V'\!\left(\frac{c}{2}\right) > \frac{V(q)}{q}.$$

If (A3) is not satisfied, it must be that $\lambda_2 = 0$, in which case the monopolist's profit is given by

$$\alpha V(q) + 2(1 - \alpha)\left(V(c) - V\!\left(\frac{c}{2}\right)\right).$$

Combining these two expressions, the monopolist's profit is given by

$$\pi(c, q) = \begin{cases} \alpha V(q) + 2(1 - \alpha)\left(V(c) - V\!\left(\frac{c}{2}\right)\right), & \text{if } V'\!\left(\frac{c}{2}\right) \leq \frac{V(q)}{q} \\[2mm] \alpha V(q) + (1 - \alpha)x(q)V'(x(q)) \\[1mm] \quad + 2(1 - \alpha)\left(V(c) - V(x(q)) - \frac{V'(x(q))}{2}(c - x(q))\right), & \text{otherwise.} \end{cases}$$

We can now prove that if $\alpha < 1$, the monopolist always offers a stock bundle for sale. For that we need to show that $\max\{\pi(c, q)\} \geq \max\{\pi^q(q)\}$. Take $\hat{q} = \arg\max\{\pi^q(q)\}$. We are going to argue that we can always find $c$ such that $\pi(c, \hat{q}) \geq \pi^q(\hat{q})$. We assume (and later verify) that $c$ is such that $V'(c/2) > V(\hat{q})/\hat{q}$. Then,

$$\pi(c, \hat{q}) - \pi^q(\hat{q}) = 2(1 - \alpha)(c - x(\hat{q}))\left(\frac{V(c) - V(x(\hat{q}))}{c - x(\hat{q})} - \frac{V'(x(\hat{q}))}{2}\right).$$

Since, $\lim_{c \to x(\hat{q})} (V(c) - V(x(\hat{q})))/(c - x(\hat{q})) = V'(x(\hat{q}))$, we can always find $\hat{c}$ close enough to $x(\hat{q})$ such that

$$\frac{V(\hat{c}) - V(x(\hat{q}))}{\hat{c} - x(\hat{q})} > \frac{V'(x(\hat{q}))}{2},$$

and, hence, $\pi(\hat{c}, \hat{q}) - \pi^q(\hat{q}) > 0$.

Finally, since $\hat{c} = x(\hat{q}) + \epsilon$ for small enough $\epsilon > 0$, $\hat{c}/2 < x(\hat{q})$ and, hence,

$$V'\left(\frac{\hat{c}}{2}\right) > V'(x(\hat{q})) = \frac{V(\hat{q})}{\hat{q}}.$$

This concludes the proof of part (i) of Theorem 5.

Next, we prove parts (iii) and (iv). The optimal policy must satisfy $V'(c) \leq V(q)/q$ and must solve

$$(c^\alpha, q^\alpha) = \arg\max_{c, q} \{\pi(c, q)\},$$

where

$$\pi(c, q) = \alpha V(q) + 2(1 - \alpha)\left(V(c) - V\left(\frac{c}{2}\right)\right)$$

$$- 2(1 - \alpha)\left(V(x(q)) - V\left(\frac{c}{2}\right) - \left(x(q) - \frac{c}{2}\right)V'(x(q))\right)\mathbf{1}\left\{V'\left(\frac{c}{2}\right) \geq \frac{V(q)}{q}\right\}.$$

Given the assumptions on $V(c)$, the function $\pi(c, q)$ is continuously differentiable in both $c$ and $q$.

If $V'(c^S/2) < V(c^*)/c^*$, the consumption induced by optimal pricing policy is $c^\alpha = c^S$ and $q^\alpha = c^*$. If the condition above is violated then either $(c^\alpha, q^\alpha)$ satisfies first-order conditions:

(i)          $$2V'(c^\alpha) = \frac{V(q^\alpha)}{q^\alpha}$$

$$\alpha q^\alpha V'(q^\alpha) = 2(1 - \alpha)(V'(q^\alpha) - V'(x(q^\alpha)))\left(\frac{c^\alpha}{2} - x(q^\alpha)\right);$$

or

(ii) $\qquad V'\left(\dfrac{c^\alpha}{2}\right) = \dfrac{V(q^\alpha)}{q^\alpha}$

$$\alpha q^\alpha V'(q^\alpha) = 2(1 - \alpha) \frac{\left(2V'(c^\alpha) - V'\left(\frac{c^\alpha}{2}\right)\right)}{\left(-V''\left(\frac{c^\alpha}{2}\right)\right)(V'(q^\alpha) - V'(x(q^\alpha)))} .$$

Clearly, in both cases

$$2V'(c^\alpha) - V'\left(\frac{c^\alpha}{2}\right) \leq 0,$$

which leads to the conclusion that $c^\alpha \geq c^S$.

Finally, we can prove part (ii) of the theorem. Let us now look at the per unit prices paid by S- and NS-consumers. The consumers without storage pay $V(q^\alpha)/q^\alpha$ for one unit of the good. The storers pay $P^\alpha/c^\alpha$, where

$$P^\alpha = \begin{cases} 2V(c^\alpha) - 2V\left(\frac{c^\alpha}{2}\right), & \text{if } V'\left(\frac{c^\alpha}{2}\right) \leq \frac{V(q^\alpha)}{q^\alpha} \\ 2V(c^\alpha) - 2V(x(q^\alpha)) + 2\left(x(q^\alpha) - \frac{c^\alpha}{2}\right)V'(x(q^\alpha)), & \text{if } V'\left(\frac{c^\alpha}{2}\right) > V'(x(q^\alpha)) = \frac{V(q^\alpha)}{q^\alpha} \end{cases} .$$

We show that in both cases $P^\alpha/c^\alpha < V(q^\alpha)/q^\alpha$. Suppose that $V'(c^\alpha/2) \leq V(q^\alpha)/q^\alpha$. Then,

$$\frac{P^\alpha}{c^\alpha} = \frac{V(c^\alpha) - V\left(\frac{c^\alpha}{2}\right)}{\frac{c^\alpha}{2}} < V'\left(\frac{c^\alpha}{2}\right) \leq \frac{V(q^\alpha)}{q^\alpha}.$$

If $V'(c^\alpha/2) > V'(x(q^\alpha)) = V(q^\alpha)/q^\alpha$, then

$$\frac{P^\alpha}{c^\alpha} = \frac{V(c^\alpha) - V(x(q^\alpha)) + \left(x(q^\alpha) - \frac{c^\alpha}{2}\right)V'(x(q^\alpha))}{\frac{c^\alpha}{2}} < V'(x(q^\alpha)) = \frac{V(q^\alpha)}{q^\alpha}.$$

This concludes the proof of Theorem 5 under the assumption that the monopolist offers a constant flow bundle $q_t = q^\alpha$.

The following lemma relaxes this assumption by showing that the monopolist can not increase profits by allowing for an arbitrary time-varying flow bundle.

LEMMA A.2: *Suppose, the monopolist can set the size of the flow bundle at each moment in time. The optimal policy is $q_t^\alpha = q^\alpha$.*

PROOF:

Take any arbitrary policy $q_t$. We show that we can find a policy with a constant size of a flow bundle, that generates weakly higher revenue than $q_t$.

The proof of this lemma consists of three steps. First, we characterize the consumer choice of consumption and purchasing sequences. Second, we show that selling an arbitrary $q_t$ generates a weakly smaller profit than selling a piece-wise constant flow bundle. And finally, we show that selling a constant flow $q^\alpha$ dominates selling any piece-wise constant flow bundle.

Clearly, the consumer will smooth his consumption up to a point, when he faces one of the two storage constraints: when either the storage is empty or it is full. Here we introduce the purely technical assumption, that the consumption as a function of time is right-continuous.

LEMMA A.3: *Suppose, the agent's consumption at time t is $x_t$. Then,*

(i) *if $V'(x_t) > V(q_t)/q_t$, the agent purchases as much flow bundles as his storage allows;*

(ii) *if $V'(x_t) < V(q_t)/q_t$, the agent does not purchase any flow bundles;*

(iii) *if $V'(x_t) = V(q_t)/q_t$, the agent is indifferent between purchasing and not purchasing flow bundle at time t.*

PROOF:

See supplementary Appendix.

Suppose, that the bulk bundle is sold at time 0, $T$, $2T$, $3T$, and so on. Let us first consider the price of a bulk bundle sold at time $T$. It is, as before, determined by no-skipping constraint. If the S-consumer skips the sale, he will purchase the flow bundle only if the per-unit price of the good is weakly lower then the marginal utility of consumption and his storage is not full (see Lemma A.3). Since we allow for multiple purchases of the flow, we can assume without loss of generality, that he is going to buy the flow at the countably many points in time. Suppose, in the interval $[0, T]$, he purchases the flow $K_T$ times at $\{t_i^T\}_{i=1}^{K_T}$. The consumption $x_t^T$ is piece-wise constant with discontinuities at $t_i^T$.

Now let us consider the bulk bundle sold at time 0. Again, if the S-consumer skips the sale at time 0, his consumption $x_t^0$ is going to be piece-wise constant with discontinuities at $t_i^0$. Let us take a coarsest refinement of the two sets of intervals that are induced by $\{t_i^T\}_{i=1}^{K_T}$ and by $\{t_i^0\}_{i=1}^{K_0}$. By $K_{0,T}$ we denote the number of intervals and by $\{t_i^{0,T}\}_{i=1}^{K_{0,T}}$ we denote their endpoints. Both $x_t^0$ and $x_t^T$ are constant within each interval. We denote their value for the interval $(t_i^{0,T}, t_{i+1}^{0,T})$ by $x_i^0$ and $x_i^T$, respectively.

By construction, for any $t \in (t_i^{0,T}, t_{i+1}^{0,T})$,

$$\frac{V(q_t)}{q_t} \geq \max\{V'(x_i^0),\ V'(x_i^T)\}.$$

Since the S-consumer is not purchasing the flow bundles within the interval $t \in \left( t_i^{0,T}, t_{i+1}^{0,T} \right)$, we can change the size of the flow bundle inside the interval $\left( t_i^{0,T}, t_{i+1}^{0,T} \right)$ to $q_i^{0,T}$. The new size of the bundle is such that the S-consumer's does not change his decisions on when and how much of the flow bundles to buy:

$$\frac{V(q_i^{0,T})}{q_i^{0,T}} = \max\{V'(x_i^0), \ V'(x_i^T)\}.$$

Note that $q_t \le q_i^{0,T}$ for any $t \in \left( t_i^{0,T}, t_{i+1}^{0,T} \right)$. This change will not affect the price of the bulk bundle and will weakly increase the profit collected from NS- consumers. From now on we will restrict our attention to the policies that are obtained by this transformation.

If the S-consumer skips a sale, he plans his consumption while having the storage filled up to its maximum capacity. He consumes from the storage when the price of the flow bundle is too high. The utility of the S-consumer in this case is

$$\int_0^{2T} \left[ V(x_t) - x_t V'(x_t) + S V'\left( \min_{\tau \in [0,2T]} \{x_\tau\} \right) \right] d_t.$$

The part of the tariff that depends on $x_i^0$ and $x_i^T$ is

$$P(\mathbf{x}^0, \mathbf{x}^T) = \alpha \sum_{i=0}^{K^{0,T}} \tau_i V(q_i^{0,T}) - (1 - \alpha)\frac{S}{2} \left( V'\left(\min_i\{x_i^0\}\right) + V'\left(\min_i\{x_i^T\}\right) \right)$$

$$- (1 - \alpha) \sum_{i=0}^{K^{0,T}} \tau_i \left[ \left( V(x_i^0) + V(x_i^T) \right) - \left( x_i^0 V'(x_i^0) + x_i^T V'(x_i^T) \right) \right],$$

where $\tau_i = t_{i+1}^{0,T} - t_i^{0,T}$ is the length of time interval $i$.

Since $q_i^{0,T} = \min\{q_i^0, q_i^T\}$, we have $V(q_i) \le \left( V(q_i^0) + V(q_i^T) \right)/2$, and, hence,

$$P(\mathbf{x}^0, \mathbf{x}^T)$$

$$\le \sum_{i=0}^{K} \tau_i \left[ \frac{\alpha}{2} V(q_i^0) - (1 - \alpha)\frac{c}{2} V'\left(\min_i\{x_i^0\}\right) - (1 - \alpha)\left( V(x_i^0) - x_i^0 V'(x_i^0) \right) \right]$$

$$+ \sum_{i=0}^{K} \tau_i \left[ \frac{\alpha}{2} V(q_i^T) - (1 - \alpha)\frac{c}{2} V'\left(\min_i\{x_i^T\}\right) - (1 - \alpha)\left( V(x_i^T) - x_i^T V'(x_i^T) \right) \right].$$

Recall, that $q_i^0$ satisfies the equation $V(q_i^0)/q_i^0 = V'(x_i^0)$, and that $q_i^T$ satisfies a similar one. Also recall that our solution $(q^\alpha, x^\alpha)$ maximizes

$$\frac{\alpha}{2} V(q(x)) - (1 - \alpha)\frac{c}{2} V'(x) - (1 - \alpha)(V(x) - x V'(x)),$$

where $q(x)$ is such that $V(q(x))/q(x) = V'(x)$. From this we conclude that $P(\mathbf{x}^0, \mathbf{x}^T)$ is lower than what we get under the flat $q^\alpha$, hence, under the optimal policy the size of the flow bundle is constant.

## REFERENCES

**Cole, Harold L., and Narayana R. Kocherlakota.** 2001. "Efficient Allocations with Hidden Income and Hidden Storage." *Review of Economic Studies* 68 (3): 523–42.

**Conlisk, John, Eitan Gerstner, and Joel Sobel.** 1984. "Cyclic Pricing by a Durable Goods Monopolist." *Quarterly Journal of Economics* 99 (3): 489–505.

**Dudine, Paolo, Igal Hendel, and Alessandro Lizzeri.** 2006. "Storable Good Monopoly: The Role of Commitment." *American Economic Review* 96 (5): 1706–19.

**Fuchs, William, and Andrzej Skrzypacz.** 2012. "Costs and Benefits of Dynamic Trading in a Lemons Market." http://cowles.econ.yale.edu/conferences/2012/sum12/et_skrzypacz.pdf.

**Hendel, Igal, Alessandro Lizzeri, and Marciano Siniscalchi.** 2005. "Efficient Sorting in a Dynamic Adverse-Selection Model." *Review of Economic Studies* 72 (2): 467–97.

**Hendel, Igal, and Aviv Nevo.** 2006a. "Measuring the Implications of Sales and Consumer Inventory Behavior." *Econometrica* 74 (6): 1637–73.

**Hendel, Igal, and Aviv Nevo.** 2006b. "Sales and consumer inventory." *RAND Journal of Economics* 37 (3): 543–61.

**Holmstrom, Bengt, and Paul Milgrom.** 1987. "Aggregation and Linearity in the Provision of Intertemporal Incentives." *Econometrica* 55 (2): 303–28.

**Hong, Pilky, R. Preston McAfee, and Ashish Nayyar.** 2002. "Equilibrium Price Dispersion with Consumer Inventories." *Journal of Economic Theory* 105 (2): 503–17.

**Jeuland, Abel P., and Chakravarthi Narasimhan.** 1985. "Dealing—Temporary Price Cuts—by Seller as a Buyer Discrimination Mechanism." *Journal of Business* 58 (3): 295–308.

**Narasimhan, Chakravarthi.** 1988. "Competitive Promotional Strategies." *Journal of Business* 61 (4): 27–49.

**Nava, Francesco, and Pasquale Schiraldi.** 2012. "Sales and Collusion in a Dynamic Market with Storage." Unpublished.

**Pesendorfer, Martin.** 2002. "Retail Sales: A Study of Pricing Behavior in Supermarkets." *Journal of Business* 75 (1): 33–66.

**Rochet, Jean-Charles, and Lars A. Stole.** 2002. "Nonlinear Pricing with Random Participation." *Review of Economic Studies* 69 (1): 277–311.

**Salop, S., and J. E. Stiglitz.** 1982. "The Theory of Sales: A Simple Model of Equilibrium Price Dispersion with Identical Agents." *American Economic Review* 72 (5): 1121–30.

**Sobel, Joel.** 1984. "The Timing of Sales." *Review of Economic Studies* 51 (3): 353–68.

**Sobel, Joel.** 1991. "Durable Goods Monopoly with Entry of New Consumers." *Econometrica* 59 (5): 1455–85.

**Stole, Lars A.** 2007. "Price Discrimination and Competition." In *Handbook of Industrial Organization, Vol. 3,* edited by Mark Armstrong and Robert Porter, 2221–99. Amsterdam: North Holland.

**Tirole, Jean.** 1988. *The Theory of Industrial Organization.* Boston: MIT Press.

**Varian, Hal R.** 1980. "A Model of Sales." *American Economic Review* 70 (4): 651–59.

**Villas-Boas, J. Miguel.** 2004. "Price Cycles in Markets with Customer Recognition." *RAND Journal of Economics* 35 (3): 486–501.

**Waldman, Michael.** 2003. "Durable Goods Theory for Real World Markets." *Journal of Economic Perspectives* 17 (1): 131–54.

**Wilson, Robert B.** 1997. *Nonlinear Pricing.* New York: Oxford University Press.