Online Appendix

Appendix B: Additional Results

In this appendix, we present additional results for our basic model. The following proposition proves how the cutoff level s changes depends on the magnitude of demand uncertainty and the degree of inter-temporal substitution in the single period problem. In Section 4.3, we have verified through numerical studies that these results extend to the case with multi-periods.

PROPOSITION 5. *In the single period problem, the cutoff level s is increasing in ρ. In particular, if M follows the 2-point distribution below:*

$$\mathcal{P}\{M=1-\kappa\}=1-q, \quad \mathcal{P}\{M=1+\kappa\}=q, \tag{EC.1}$$

where $q=\frac{1}{2}$ and $\kappa\leq 1$, then s is decreasing in κ.

Proof: By Theorem 1, $z=z^*(x)$ solves the problem

$$\max_{0\leq z\leq x}\pi(\tau,z)=(\beta+\rho)\theta(\tau)+p\mathbb{E}(z\wedge\alpha M)-\rho\theta(\tau\wedge z),$$

where $\theta(\tau\wedge z)$ is increasing in z since τ maximizes the concave function $\theta(x)$. Thus, the objective function is supermodular in $(-\rho,z,x)$, implying that $z^*(x)$ is increasing in $-\rho$ and x by Theorem 2.8.2 in Topkis (1998). This ensures that $s=\sup\{x\geq 0: z^*(x)=0\}$ is increasing in ρ.

To show the monotonicity in κ, observe from the expression of $\pi(\tau,z)$ that the cutoff level s is the (larger) root of $f(z)=0$ for

$$f(z)=p\mathbb{E}(z\wedge\alpha M)-\rho\left[r\mathbb{E}(z\wedge\tau\wedge\alpha M)-c(z\wedge\tau)\right].$$

When M follows the given 2-point distribution, $f(z)$ consists of three linear pieces. In particular,

$$f(z)=\begin{cases}[p-\rho(r-c)]z, & \text{if } z\leq\alpha(1-\kappa)\\ \alpha p\mathbb{E}M-\rho[r\mathbb{E}(\tau\wedge\alpha M)-c\tau], & \text{if } z\geq\alpha(1+\kappa).\end{cases}$$

and if $\alpha(1-\kappa)\leq z\leq\alpha(1+\kappa)$, then

$$\begin{aligned}f(z)&=\begin{cases}p\mathbb{E}(z\wedge\alpha M)-\rho(r-c)\alpha(1-\kappa) & \text{if } \tau=\alpha(1-\kappa)\\ (p-\rho r)\mathbb{E}(z\wedge\alpha M)+\rho cz, & \text{if } \tau=\alpha(1+\kappa)\end{cases}\\ &=\begin{cases}p[(1-q)\alpha(1-\kappa)+qz]-\rho(r-c)\alpha(1-\kappa) & \text{if } \ qr\leq c\\ (p-\rho r)[(1-q)\alpha(1-\kappa)+qz]+\rho cz, & \text{if } \ qr\geq c\end{cases}\end{aligned}$$

To find the cutoff s, consider the following cases:

- $c \geq qr$: $\tau = \alpha(1-\kappa)$. If $p \geq \rho(r-c)$, $f(z)$ always increases in z, and thus $s = 0$. If $\rho(r-c)(1-\kappa) < p < \rho(r-c)$, $f(z)$ decreases and then increases in z, attaining a positive value at $z = \alpha(1+\kappa)$. Hence, $s = \left[\frac{\rho(r-c)-p(1-q)}{qp}\right]\alpha(1-\kappa)$. If $p \leq \rho(r-c)(1-\kappa)$, $f(z)$ first decreases and then increases in z, attaining a nonpositive value at $z = \alpha(1+\kappa)$. Hence, $s = +\infty$.
- $c < qr$: $\tau = \alpha(1+\kappa)$. If $p \geq \rho(r-c)$, $f(z)$ always increases in z, and thus $s = 0$. If $\rho(r-c(1+\kappa)) < p < \rho(r-c)$, $f(z)$ decreases and then increases in z, attaining a positive value at $z = \alpha(1+\kappa)$. Hence, $s = \left[\frac{(\rho r-p)(1-q)}{\rho c - q(\rho r - p)}\right]\alpha(1-\kappa)$. If $p \leq \rho(r-c(1+\kappa))$, $f(z)$ first decreases and then may increase in z, attaining a nonpositive value at $z = \alpha(1+\kappa)$. Hence, $s = +\infty$.

In both cases above, s decreases in κ. □

Finally, we present a result about the trajectory of system (2). In particular, given a series of independent and identical distributed random variables $\{M_n : n \geq 1\}$ with $M_n \leq M_{\max}$, we call the sequence $\{x_n : n \geq 0\}$ a corresponding trajectory of system (2) if for each $n \geq 0$,

$$x_{n+1} = [D(y^*(x_n), z^*(x_n)) - D(\alpha M, z^*(x_n))]^+,$$

where x_0 is predetermined. Obviously $\{x_n : n \geq 1\}$ forms a homogeneous Markovian process. Furthermore, it can be shown that its state space belongs to a compact set: as shown in the proof of Theorem 3, $y^*(x_n) \leq \alpha M_{\max}$, implying $x_{n+1} \leq D(y^*(x_n), z^*(x_n)) \leq (\beta+\rho)y^*(x_n) \leq (\beta+\rho)\alpha M_{\max}$. Hence, this homogeneous Markovian process has a stationary distribution. That is, there exists a probability distribution Φ such that if the state x_n follows Φ on the left side of problem (2), then the state x_{n+1} in next period under the optimal decisions follows the same distribution Φ. We summarize the result as follows:

PROPOSITION 6. *Any trajectory of system* (2) *forms a homogeneous Markovian process which admits a stationary distribution.*

Appendix C: Extensions

C1. Consumer Model without Outside Option in the Second Phase

Here, we present the micro-consumer model mentioned in Section 5.1.2 in greater detail. In particular, we describe a rational expectation framework in which consumers make their decisions based on their correct beliefs regarding the fill rate in the second phase, and the firm makes markdown and inventory decision based on its correct belief regarding purchasing behavior of customers. Note that the second-phase fill rate λ_2 is determined by the second-phase demand $(1-\alpha)M + \rho(\alpha M - z)^+$ and the inventory y^0, both of which depend on the initial inventory x. Nevertheless, in practice consumers can hardly observe or access to information about markdown ration z, inventory y^0, or initial inventory x, as these are the firm's internal operating data. Thus, the consumers need to form a rational expectation about the second-phase fill rate. Assume that pricing is exogenous and that the system starts from the steady state characterized in Proposition 6 in Appendix B. Denote the steady state by X_0 and the stationary distribution by Φ. In such a case, the second-phase fill rate λ_2 is stationary over time. Denote by $\tilde{\lambda}_2$ the second-phase fill rate anticipated by the consumers. In the meantime, the firm also conjectures about the consumers' purchasing choices, characterized by the two fractions α and ρ. Let $\tilde{\alpha}$ and $\tilde{\rho}$ be the firm's beliefs.

Definition 1. A rational-expectation equilibrium is sustained when the following conditions are satisfied (Su and Zhang 2008; Su 2010):

(a) For each realization, x, of the initial state X_0, the firm maximizes expected profit subject to beliefs about consumer behavior: $(y^{0*}(x), z^*(x))$ is the solution to the following problem:

$$v(x) = \max_{y^0, z} \left\{ \pi^0(y^0, z) + \gamma \mathbb{E} v(\tilde{x}) : y^0 \geq 0, 0 \leq z \leq x \right\}, \tag{EC.2}$$

where $\tilde{x} = \left[y^0 - (1-\tilde{\alpha})M - \tilde{\rho}\,(\tilde{\alpha}M - z)^+ \right]^+$ and

$$\pi^0(y^0, z) = p\mathbb{E}\left[z \wedge (\tilde{\alpha}M)\right] + r\mathbb{E}\left\{ y^0 \wedge \left[(1-\tilde{\alpha})M + \tilde{\rho}\,(\tilde{\alpha}M - z)^+ \right] \right\} - cy^0.$$

Under the optimal solution, the expected second-phase fill rate is

$$\lambda_2^* = \mathbb{E}_{X_0, M} \left\{ 1 \wedge \frac{y^{0*}(X_0)}{(1-\tilde{\alpha})M + \tilde{\rho}[\tilde{\alpha}M - z^*(X_0)]^+} \right\}.$$

(b) The consumers purchase in the phase which maximizes their expected utility subject to their belief about the second-phase fill rate: $\alpha^* = \alpha^*(\tilde{\lambda}_2)$, $\rho^* = \rho^*(\tilde{\lambda}_2)$.

(c) The firm's beliefs are rational: $\alpha^* = \tilde{\alpha}$, $\rho^* = \tilde{\rho}$.

(d) The consumers' beliefs are rational: $\lambda_2^* = \tilde{\lambda}_2$.

A full characterization of the rational-expectation equilibrium is difficult, even numerically. Nevertheless, since the firm's markdown and inventory decision is optimized for given beliefs about consumer purchasing behavior, the firm's best response, $[y^{0*}(x), z^*(x)]$, to a given pair $(\tilde{\alpha}, \tilde{\rho})$ is exactly the same as that in our base model. That is, if the aggregate demand follows a 2-point distribution, then a bang-bang solution is optimal. The proposition below mirrors Proposition 4.

Proposition 7. *Assume that the aggregate demand M follows a 2-point distribution. Whenever a rational-expectation equilibrium is sustained, a bang-bang solution for the markdown decision is optimal, i.e., z^* equals to either x or 0.*

C2. Endogenous Markdown Pricing

We focus on the single-period problem below:

$$v(x) = \max_{y,z,p} \{\pi(y,z,p) : y \geq 0, z \in [0,x], p \in \mathcal{P}\}, \qquad \text{(EC.3)}$$

where

$$\begin{aligned}
\pi(y,z,p) &= p\mathbb{E}\left\{z \wedge \left[d_1(p)M^0\right]\right\} - cD(y,z,p) + r\mathbb{E}D(y \wedge M^0, z, p),\\
D(y,z,p) &= d_2(p)y + \{d_3(p)y - [d_3(p)/d_1(p)]z\}^+,\\
\mathcal{P} &= \{p_1, p_2, ..., p_m\},
\end{aligned}$$

and recall the definitions of $d_1(p)$, $d_2(p)$, and $d_3(p)$:

1. $d_1(p) = \left(\frac{r-p}{1-\delta} \wedge 1\right) - \frac{p}{\delta}$: the portion of consumers who attempt to make purchases in the first phase;
2. $d_2(p) = 1 - \left(\frac{r-p}{1-\delta} \wedge 1\right)$: the portion of consumers who have high valuation that they disdain to buy in the first phase and only buy the fresh product in the second phase; and
3. $d_3(p) = \left(\frac{r-p}{1-\delta} \wedge 1\right) - r$: the portion of consumers who have a valuation higher than r but they are attracted by the promotion and try to make purchases in the first phase. If rationed in the first phase, this group of consumers will attempt to make purchase in the second phase.

Assume that $p \in \mathcal{P} = \{p_1, p_2, ..., p_m\}$ with $0 < p_1 < \cdots < p_m < \delta r$. To preclude the unrealistic case where the markdown price is extremely low that markdown sales completely cannibalize regular sales, we assume that $d_2(p) > 0$ for all $p \in \mathcal{P}$. Under these conditions, the following proposition proves the optimality of the bang-bang policy and also characterizes the cutoff level.

PROPOSITION 8. *Consider the single-period problem* (EC.3) *and assume that* $p \in \mathcal{P} = \{p_1, p_2, ..., p_m\}$ *where* $0 < p_1 < \cdots < p_m < \delta r$ *and* $d_2(p) > 0$ *for all* $p \in \mathcal{P}$*. Then* $y^*(x) = \tau^0$*, where* τ^0 *maximizes the function* $\theta^0(x) = r\mathbb{E}(x \wedge M^0) - cx$*. Furthermore, there exists a cutoff level* s *such that* $z^*(x) = 0$ *if* $x < s$*, and* $z^*(x) = x$ *otherwise. In particular, if* M^0 *follows the* 2*-point distribution in* (EC.1)*, then* s *is decreasing in* κ*.*

C3. An Extended Consumer Model with Consumers' Shopping Cost

In this extended model we assume that a consumer incurs a cost $\phi > 0$ for each visit to the store. Consider the case where consumers do not have outside purchasing option and the firm's prices are exogenously given. A consumer with valuation v attempts to buy in the first phase if and only if

$$\lambda_1(\delta v - p) + (1-\lambda_1)[\lambda_2(v-r) - \phi]^+ - \phi \geq [\lambda_2(v-r) - \phi]^+, \qquad \text{(EC.4)}$$

where $[\lambda_2(v-r)-\phi]^+$ is the expected payoff of either voluntarily skipping the first phase or being rationing out in the first phase: in either case, the consumer can choose to either visit the store again in the second phase, paying the inconvenience cost ϕ and getting the product with probability λ_2, or not come back to the second phase and get zero surplus. Clearly, a customer who is rationed out in the first phase will come back to the second phase if and only if $\lambda_2(v-r)-\phi \geq 0$, or $v \geq \frac{\phi}{\lambda_2}+r$. It is then clear that a consumer is less likely to come back to the second phase when ϕ increases. Also, from (EC.4), it is straightforward to show that a consumer is less likely to buy in the first phase when ϕ increases or when λ_1 decreases.

In the meanwhile, a consumer voluntarily skips the first phase and attempts to buy in the second phase if and only if

$$\lambda_2(v-r)-\phi \geq \{\lambda_1(\delta v-p)+(1-\lambda_1)[\lambda_2(v-r)-\phi]^+-\phi\}^+. \tag{EC.5}$$

Also, a consumer does not make the attempt to buy in either phase if and only if

$$0 \geq \left\{\lambda_1(\delta v-p)+(1-\lambda_1)[\lambda_2(v-r)-\phi]^+-\phi\right\} \vee [\lambda_2(v-r)-\phi], \tag{EC.6}$$

which implies $v \leq \left[\frac{1}{\delta}(p+\frac{\phi}{\lambda_1})\right] \wedge \left(r+\frac{\phi}{\lambda_2}\right)$. To preclude trivial situations where no consumer buys in either phase due to a prohibitively high shopping cost, we assume that ϕ is sufficiently low such that the equilibrium $(\alpha^*, \rho^*, \lambda_1^*, \lambda_2^*)$ satisfies

$$\left[\frac{1}{\delta}(p+\frac{\phi}{\lambda_1^*})\right] \wedge \left(r+\frac{\phi}{\lambda_2^*}\right) < 1. \tag{EC.7}$$

Note that the condition (EC.7) is ensured when ϕ equals to zero. Thus, by continuity of the problem, there exists a range of low ϕ's under which the condition (EC.7) is satisfied. Also note that, when $\phi > 0$, condition (EC.7) implies either $\lambda_1^* > 0$ or $\lambda_2^* > 0$ or both.

The following proposition characterizes consumers' purchasing behavior and resultant market segmentation, for given belief about the fill rates in the two phases.

PROPOSITION 9. *Given λ_1 and λ_2 satisfying $\left[\frac{1}{\delta}\left(p+\frac{\phi}{\lambda_1}\right)\right] \wedge \left[r+\frac{\phi}{\lambda_2}\right] < 1$, the consumers' purchasing behavior under shopping cost ϕ is as follows:*

(i) If $\lambda_2 \leq \delta$ and $p \leq \delta r + \phi(\frac{\delta}{\lambda_2}-\frac{1}{\lambda_1})$, consumers with $v \geq \frac{1}{\delta}[p+\frac{\phi}{\lambda_1}]$ buy in the first phase and those with $v < \frac{1}{\delta}[p+\frac{\phi}{\lambda_1}]$ do not buy in either phase, i.e., $\alpha^*(\lambda_1,\lambda_2)=1$ and $\rho^*(\lambda_1,\lambda_2)=\frac{\left[1-(r+\frac{\phi}{\lambda_2})\right]^+}{1-\frac{1}{\delta}(p+\frac{\phi}{\lambda_1})}$.

(ii) If $\lambda_2 \leq \delta$ and $p > \delta r + \phi(\frac{\delta}{\lambda_2}-\frac{1}{\lambda_1})$, consumers with $v \geq \frac{1}{\delta-\lambda_2}(p+\frac{\phi}{\lambda_1}-\phi-\lambda_2 r)$ buy in the first phase, those with $v \in [r+\frac{\phi}{\lambda_2}, \frac{1}{\delta-\lambda_2}(p+\frac{\phi}{\lambda_1}-\phi-\lambda_2 r)]$ buy in the second phase, and those with $v \leq r+\frac{\phi}{\lambda_2}$ do not buy in either phase, i.e., $\alpha^*(\lambda_1,\lambda_2)=\frac{\left[1-\frac{1}{\delta-\lambda_2}(p+\frac{\phi}{\lambda_1}-\phi-\lambda_2 r)\right]^+}{1-(r+\frac{\phi}{\lambda_2})}$ and $\rho^*(\lambda_1,\lambda_2)=1$.

(iii) If $\lambda_2 > \delta$ and $p \le \delta r + \phi(\frac{\delta}{\lambda_2} - \frac{1}{\lambda_1})$, consumers with $v \in [\frac{1}{\delta}(p + \frac{\phi}{\lambda_1}), \frac{1}{\delta-\lambda_2}(p + \frac{\phi}{\lambda_1} - \phi - \lambda_2 r)]$ buy in the first phase, those with $v > \frac{1}{\delta-\lambda_2}(p + \frac{\phi}{\lambda_1} - \phi - \lambda_2 r)$ buy in the second phase and those with $v < \frac{1}{\delta}(p + \frac{\phi}{\lambda_1})$ do not buy in either phase, i.e.,

$$\alpha^*(\lambda_1,\lambda_2) = \frac{1\wedge\left[\frac{1}{\delta-\lambda_2}(p+\frac{\phi}{\lambda_1}-\phi-\lambda_2 r)\right]-\frac{1}{\delta}(p+\frac{\phi}{\lambda_1})}{1-\frac{1}{\delta}(p+\frac{\phi}{\lambda_1})} \text{ and } \rho^*(\lambda_1,\lambda_2) = \frac{\left\{1\wedge\left[\frac{1}{\delta-\lambda_2}(p+\frac{\phi}{\lambda_1}-\phi-\lambda_2 r)\right]-(r+\frac{\phi}{\lambda_2})\right\}^+}{1\wedge\left[\frac{1}{\delta-\lambda_2}(p+\frac{\phi}{\lambda_1}-\phi-\lambda_2 r)\right]-\frac{1}{\delta}(p+\frac{\phi}{\lambda_1})}.$$

(iv) If $\lambda_2 > \delta$ and $p > \delta r + \phi(\frac{\delta}{\lambda_2} - \frac{1}{\lambda_1})$, consumers with $v \ge r + \frac{\phi}{\lambda_2}$ buy in the second phase and those with $v < r + \frac{\phi}{\lambda_2}$ do not buy in either phase, i.e., $\alpha^*(\lambda_1,\lambda_2) = 0$, and $\rho^*(\lambda_1,\lambda_2)$ is arbitrary.

Particularly, in all the cases above, if $\alpha^*(\lambda_1,\lambda_2) = 0$, then $\rho^*(\lambda_1,\lambda_2)$ is arbitrary.

Similar to Definition 1, we now define a rational-expectation equilibrium for this extended consumer model.

Definition 2. A rational-expectation equilibrium is sustained when the following conditions are satisfied (Su and Zhang 2008; Su 2010):

1. For each realization, x, of the initial state X_0, the firm maximizes expected profit subject to beliefs about consumer behavior $(\tilde{\alpha}, \tilde{\rho})$. Specifically, $[y^{0*}(x), z^*(x)]$ is the solution to the following problem:

$$v(x) = \max_{y^0,z} \left\{\pi^0(y^0,z) + \gamma\mathbb{E}v(\tilde{x}) : y^0 \ge 0, 0 \le z \le x\right\}, \tag{EC.8}$$

where

$$\pi^0(y^0,z) = p\mathbb{E}\left[z \wedge (\tilde{\alpha}M)\right] + r\mathbb{E}\left\{y^0 \wedge \left[(1-\tilde{\alpha})M + \tilde{\rho}\left(\tilde{\alpha}M - z\right)^+\right]\right\} - cy^0,$$
$$\tilde{x} = \left[y^0 - (1-\tilde{\alpha})M - \tilde{\rho}\left(\tilde{\alpha}M - z\right)^+\right]^+,$$

and M is identically distributed as the aggregate demand M_n. Under the optimal solution, the expected fill rates in the two phases are

$$\lambda_1^* = \mathbb{E}_{X_0,M}\left[1 \wedge \frac{z^*(X_0)}{\tilde{\alpha}M}\right] \text{ and } \lambda_2^* = \mathbb{E}_{X_0,M}\left[1 \wedge \frac{y^{0*}(X_0)}{(1-\tilde{\alpha})M + \tilde{\rho}(\tilde{\alpha}M - z^*(X_0))^+}\right],$$

respectively.

2. The consumers purchase in the phase which maximizes their expected utility subject to their belief about the fill rates $(\tilde{\lambda}_1, \tilde{\lambda}_2)$: $\alpha^* = \alpha^*(\tilde{\lambda}_1, \tilde{\lambda}_2)$, $\rho^* = \rho^*(\tilde{\lambda}_1, \tilde{\lambda}_2)$, where $\alpha^*(\cdot,\cdot)$ and $\rho^*(\cdot,\cdot)$ are as defined in Proposition 9.
3. The firm's beliefs are rational: $\alpha^* = \tilde{\alpha}$, $\rho^* = \tilde{\rho}$.
4. The consumers' beliefs are rational: $\lambda_1^* = \tilde{\lambda}_1$, $\lambda_2^* = \tilde{\lambda}_2$.

Similar to Proposition 7, since the firm's markdown and inventory decision is optimized for given beliefs about consumer purchasing behavior, the firm's best response, $[y^{0*}(x), z^*(x)]$, to a given pair $(\tilde{\alpha}, \tilde{\rho})$ is exactly the same as that in our base model. Hence, if the aggregate demand follows a 2-point distribution, then a bang-bang solution is optimal.

To further characterize the effects of the shopping cost ϕ, we solve for the rational-expectation equilibrium in a special case. In preparation, we first characterize the firm's optimal cutoff level for its markdown decision in the special case, as below.

PROPOSITION 10. *Consider the single-period problem (i.e., $\gamma = 0$) where the aggregate demand M follows a two-point distribution: $\mathcal{P}(M = M_{\min}) = 1 - q$ and $\mathcal{P}(M = M_{\max}) = q$. For given α and ρ, the firm's optimal strategy $[y^{0*}(x), z^*(x)]$ is as below:*

$$y^{0*}(x) = \beta\tau + \rho[\tau - z^*(x)]^+, \quad z^*(x) = \begin{cases} x, & \text{if } x \geq s \\ 0, & \text{if } x < s. \end{cases}$$

where $\beta = \frac{1-\alpha}{\alpha}$, τ maximizes $r\mathbb{E}(x \wedge \alpha M) - cx$, and the cutoff s is as follows:

(a) Suppose $c \geq qr$. $s = 0$ if $p \geq \rho(r - c)$, $s = \left[\frac{\rho(r-c) - p(1-q)}{qp}\right] \alpha M_{\min}$ if $\rho(r - c)\frac{M_{\min}}{\mathbb{E}(M)} < p < \rho(r - c)$, and $s = M_{\max}$ if $p \leq \rho(r - c)\frac{M_{\min}}{\mathbb{E}(M)}$.

(b) Suppose $c < qr$. $s = 0$ if $p \geq \rho(r - c)$, $s = \left[\frac{(\rho r - p)(1-q)}{\rho c - q(\rho r - p)}\right] \alpha M_{\min}$ if $\rho(r - c\frac{M_{\max}}{\mathbb{E}(M)}) < p < \rho(r - c)$, and $s = M_{\max}$ if $p \leq \rho(r - c\frac{M_{\max}}{\mathbb{E}(M)})$.

Assuming that the firm and consumers share the same prior belief about the initial state X_0: X_0 is uniformly distributed between zero and $M_{\max}$, Proposition 10 allows us to compute the expected fill rates in both phases when the firm follows its optimal inventory and markdown decisions in response to given beliefs about α and ρ. Corollary 1 follows.

COROLLARY 1. *Consider the single-period problem (i.e., $\gamma = 0$) where M follows a two-point distribution: $\mathcal{P}(M = M_{\min}) = 1 - q$ and $\mathcal{P}(M = M_{\max}) = q$. Assume $c < qr$. Also, assume that the firm and consumers share the same prior belief about the initial state X_0: X_0 is uniformly distributed between zero and $M_{\max}$. For given α and ρ, $\lambda_2^* = 1$ and the expression of λ_1^* is as follows:*

(a) if $p \geq \rho(r - c)$, $\lambda_1^* = 1 - \alpha\frac{\mathbb{E}(M)}{2M_{\max}}$;

(b) if $\rho(r - c\frac{M_{\max}}{\mathbb{E}(M)}) < p < \rho(r - c)$, $\lambda_1^* = 1 - \alpha\left[\frac{(\rho r - p)(1-q)}{\rho c - q(\rho r - p)}\frac{M_{\min}}{M_{\max}} + \frac{q}{2}\left(1 - \frac{(\rho r - p)(1-q)}{\rho c - q(\rho r - p)}\frac{M_{\min}}{M_{\max}}\right)^2\right]$;

(c) if $p \leq \rho(r - c\frac{M_{\max}}{\mathbb{E}(M)})$, $\lambda_1^* = 0$.

Below we focus on the case defined in Corollary 1, that is, $\gamma = 0$ and M follows a two-point distribution: $\mathcal{P}(M = M_{\min}) = 1 - q$ and $\mathcal{P}(M = M_{\max}) = q$. Assume $c < qr$. Also, assume that the firm and consumers share the same prior belief about the initial state X_0: X_0 is uniformly distributed between zero and $M_{\max}$. In addition, we make the following further assumptions on the parameters:

$$\text{(A.1)} \quad \delta < 1$$
$$\text{(A.2)} \quad r + \phi < 1$$
$$\text{(A.3)} \quad p < \delta r - \phi(1-\delta)$$
$$\text{(A.4)} \quad r + \delta \le 1$$

(A.1) reflects the fact that the quality of the product deteriorates over time. (A.2) requires that phase-2 demand is positive if the firm sells only in phase 2 (i.e., $z = 0$) and also guarantees fulfillment of all demand in phase 2 (i.e., $\lambda_2 = 1$). If (A.3) is violated, phase-1 demand is zero even if the fill rates in both phases are one (i.e., $\lambda_1 = \lambda_2 = 1$). Thus, (A.3) sets an upper bound for the discount price below which the markdown sales are attractive to some consumers, when product availability is guaranteed in both phases. Assumption (A.4) is for analytical convenience and can be relaxed without affecting the results qualitatively.

Now we derive the rational expectation under these conditions. First note that ($\lambda_1^* = 0$, $\lambda_2^* = 1$, $\alpha^* = 0$, ρ arbitrary) is always an equilibrium. We then focus on searching for equilibrium with $\lambda_1^* > 0$. Also, to focus on the most realistic case, we search only the equilibrium with $\alpha^* < 1$, i.e., the markdown sales does not completely cannibalize the regular sales.

Proposition 11. *Under the conditions in Corollary 1 and Assumptions (A.1) through (A.4), there exists a unique rational-expectation equilibrium where $\lambda_1^* > 0$ and $\alpha^* < 1$. The equilibrium profile is as follows:*

$$\alpha^* = \frac{1-\lambda_1^*}{W}, \quad \rho^* = \delta, \quad \lambda_1^* \text{ is a solution to } g(\lambda_1) = 0, \quad \lambda_2^* = 1,$$

where $g(\lambda_1) = (1-\delta)(\delta-p)\lambda_1^2 + (W(r\delta - p + \delta\phi) + (\delta-1)(\delta-p+\phi))\lambda_1 - (W\phi + \phi(\delta-1))$, $W = \frac{E(M)}{2M_{\max}}$ *if* $p \ge \delta(r-c)$ *and* $W = \frac{(\delta r-p)(1-q)}{\delta c - q(\delta r - p)}\frac{M_{\min}}{M_{\max}} + \frac{q}{2}(1 - \frac{(\delta r-p)(1-q)}{\delta c - q(\delta r-p)}\frac{M_{\min}}{M_{\max}})^2$ *if* $\delta(r - c\frac{M_{\max}}{\mathbb{E}(M)}) < p < \delta(r-c)$. *Furthermore,* λ_1^* *increases in* ϕ *and* α^* *decreases in* ϕ*, implying that the cutoff level* s *in equilibrium also decreases in* ϕ.

As shown in Proposition 11, a higher shopping cost disincentives strategic purchasing behavior and makes it more likely for the firm to sell during markdown phase.

C4. An Extended Model with Additional Costs for the Firm

Here, we first present a model where there is a holding cost for inventories carried from one period to the next. We show that the model is reduced to the base model by proper redefinitions of terms.

Because $\tilde{x} = D(y \vee \alpha M, z) - D(\alpha M, z) = D(y, z) - D(y \wedge \alpha M, z)$, the objective function of problem (2) can be reformulated as

$$g(y,z) = p\mathbb{E}\left[z \wedge (\alpha M)\right] + (r-c)D(y,z) - r\mathbb{E}\tilde{x} + \gamma \mathbb{E}v(\tilde{x}).$$

Now assume that the firm incurs a holding cost h for each unit carried across two periods. In this case, the corresponding objective function becomes

$$\begin{aligned} g_h(y,z) &= p\mathbb{E}\left[z \wedge (\alpha M)\right] + (r-c)D(y,z) - (r+h)\mathbb{E}\tilde{x} + \gamma \mathbb{E}v(\tilde{x}) \\ &= p\mathbb{E}\left[z \wedge (\alpha M)\right] + \left[(r+h) - (c+h)\right]D(y,z) - (r+h)\mathbb{E}\tilde{x} + \gamma \mathbb{E}v(\tilde{x}). \end{aligned}$$

It is exactly the same as $g(y,z)$ with r replaced by $(r+h)$ and c replaced by $(c+h)$, respectively. Hence, all the results remain the same as in the base model.

In addition, consider the case when the firm incurs a fixed ordering cost. Due to the perishability of the product, assuming that the store is always in business, it needs to reorder in every period and thus adding a fixed ordering cost does not change our results.

Appendix D: Proofs of Appendix C

Proof of Proposition 8

In preparation, define $\theta^0(x) := r\mathbb{E}[x \wedge M^0] - cx$ and let τ^0 be its (smallest) maximizer. We start from deriving some important properties of the single-period profit function. For markdown price p satisfying $d_2(p) > 0$, or equivalently, $\frac{r-p}{1-\delta} < 1$, it can be verified that $d_3(p) = \delta d_1(p) = \frac{\delta r - p}{1-\delta}$ and $D(y,z,p) = (1-r)y - \delta\left\{[d_1(p)y] \wedge z\right\}$.By substituting the expressions of $D(y,z,p)$ and $D(y \wedge M^0, z, p)$, we have:

$$\pi(y,z,p) = p\mathbb{E}\left\{z \wedge [d_1(p)M^0]\right\} + (1-r)\theta^0(y) - [\delta d_1(p)]\,\theta^0\left(\tfrac{z}{d_1(p)} \wedge y\right),$$

which depends on y via the term $(1-r)\theta^0(y)$ if $z < yd_1(p)$ and the term $[1 - r - \delta d_1(p)]\theta^0(y)$ otherwise. Since $1 - r - \delta d_1(p) = d_2(p) > 0$, similar to the proof of Theorem 1, we can prove that $y^*(x) = \tau^0$. Furthermore, if let $\Pi(z,p) = \pi(\tau^0, z, p) - (1-r)\theta^0(\tau^0)$, i.e.,

$$\begin{aligned} \Pi(z,p) &= p\mathbb{E}\left\{z \wedge [d_1(p)M^0]\right\} - \delta r\mathbb{E}\left\{z \wedge [d_1(p)\tau^0] \wedge [d_1(p)M^0]\right\} + \delta c\left\{z \wedge [d_1(p)\tau^0]\right\} \\ &= \begin{cases} (p-\delta r)\,\mathbb{E}\left\{z \wedge [d_1(p)M^0]\right\} + \delta cz, & \text{if } z \le d_1(p)\tau^0, \\ p\mathbb{E}\left\{z \wedge [d_1(p)M^0]\right\} - \delta d_1(p)\theta^0(\tau^0), & \text{if } z \ge d_1(p)\tau^0, \end{cases} \end{aligned} \tag{EC.9}$$

then it is straightforward to see that $[z^*(x), p^*(x)]$ solves the problem

$$\max_{z,p}\left\{\Pi(z,p) : z \in [0,x], p \in \mathcal{P}\right\}.$$

Several observations are made on $\Pi(z,p)$ for any $z \ge 0$ and $p \in (0, \delta r)$ as follows.

1. When $z \le d_1(p)\tau^0$, $\Pi(z,p)$ is convex in z by $p < \delta r$; moreover, it is increasing in p since $(p-\delta r)d_1(p) = \frac{(p-\delta r)^2}{-\delta(1-\delta)}$ is increasing in p when $0 < p < \delta r$.
2. When $z \ge d_1(p)\tau^0$, $\Pi(z,p)$ is increasing in z by $p \ge 0$; moreover, it is concave in p because $d_1(p)$ is linearly decreasing in p.

In summary, $\Pi(z,p)$ is increasing in p if $z \le d_1(p)\tau^0$ and concave in p otherwise. This ensures that $d_1(p^*(x))\tau^0 \le z^*(x)$ and the quasi-concavity of $\Pi(z,p)$ in $p \in \mathcal{P}$ for any $z \ge 0$. On the other hand, $\Pi(z,p)$ is convex in z when $z \le d_1(p)\tau^0$ and increasing in z otherwise. This ensures the quasi-convexity of $\Pi(z,p)$ in $z \ge 0$ for any $p \in \mathcal{P}$, implying that $z^*(x) \in \{0, x\}$.

We now distinguish whether $z^*(x) = x$ or $z^*(x) = 0$. By $z^*(x) \in \{0, x\}$ and $\Pi(0,p) = 0$ for any p,

$$\max_{z,p} \{\Pi(z,p) : z \in [0,x], p \in \mathcal{P}\} = 0 \vee \max_p \{\Pi(x,p) : p \in \mathcal{P}\}.$$

Thus, $z^*(x) = 0$ if and only if $\max_p \{\Pi(x,p) : p \in \mathcal{P}\} \le 0$ or equivalently,

$$\forall p \in \mathcal{P}: \quad x < s(p) \triangleq \inf\{x > 0 : \Pi(x,p) > 0\}.$$

Note that we specify $s(p) = +\infty$ if the right side of its definition admits an empty set, i.e., $\Pi(x,p) \le 0$ for all $x > 0$. Moreover, if $s(p) < +\infty$, its continuity immediately follows from the continuity of $\Pi(x,p)$. In the following we shall prove that $s(p)$ is decreasing in p. With such result, $x < s(p)$ for all $p \in \mathcal{P}$ if and only if $x < s$ with $s = s(p_m)$, i.e., $z^*(x) = 0$ if $x < s$ and $z^*(x) = x$ if $x > s$.

To characterize $s(p)$, defined $P_{\min} = \delta\theta^0(\tau^0)/(\mathbb{E}M^0)$ and $P_{\max} = \delta(r-c)$.

1. If $p = \delta r$, then obviously $\pi(x, \delta r) = 0$ for all $x > 0$, implying that $s(p) = 0$.
2. If $P_{\max} \le p < \delta r$, then $\tau^0 \ge M_{\min} > 0$ and $d_1(p)M^0 > d_1(\delta r)M_{\min} = 0$. By (EC.9),

$$\partial_z^+ \Pi(0,p) \triangleq \lim_{z \downarrow 0} \tfrac{1}{z}\Pi(z,p) = (p - P_{\max}) + \lim_{z\downarrow 0}(p-\delta r)\,\mathbb{E}\left\{0 \wedge \left[\tfrac{1}{z}d_1(p)M^0 - 1\right]\right\} = p - P_{\max},$$

   where the last equality is ensured by *dominate control theory*. By its convexity in z when $z \le d_1(p)\tau^0$, $\Pi(z,p)$ is increasing in $z \in [0, d_1(p)\tau^0]$. Because $\Pi(z,p)$ is also increasing in z when $z \ge d_1(p)\tau^0$, $\Pi(x,p) \ge \Pi(0,p) = 0$ for any $x > 0$, implying $s(p) = 0$.
3. When $p \le P_{\min}$, by (EC.9), we know that

$$\Pi(+\infty, p) \triangleq \lim_{z\to\infty} \Pi(z,p) = \left[d_1(p)(\mathbb{E}M^0)\right](p - P_{\min}) \le 0.$$

   By $\Pi(0,p) = 0$ and its quasi-convexity in z, $\Pi(x,p) \le 0$ for all $x \ge 0$, implying $s(p) = +\infty$.
4. When $P_{\min} < p < P_{\max}$, we can verify that $\partial_z^+\Pi(0,p) < 0 < \Pi(+\infty, p)$ similar to the previous two steps. It together with the quasi-convexity of $\Pi(z,p)$ in z suggests that as z increases, $\Pi(z,p)$ is firstly strictly decreasing from $\Pi(0,p) = 0$, and then increasing up to some positive value. Thus, $s(p)$ is finite and positive.

The above discussion shows that $s(p)=0$ if $p\geq P_{\max}$, $s(p)$ is finite and positive if $P_{\max}>p>P_{\min}$, and $s=+\infty$ if $p\leq P_{\min}$. Next we show $s(p)$ is decreasing in $p\in[P_{\min},P_{\max}]$. Observe that

$$\widetilde{s}(p)\triangleq[s(p)]/[d_1(p)]=\min\{x\geq 0:\Pi(xd_1(p),p)\geq 0\}=\min\{x\geq 0:\widetilde{\Pi}(x,p)\geq 0\},$$

where $\widetilde{\Pi}(x,p)=[\delta d_1(p)]^{-1}\Pi(xd_1(p),p)$, i.e., $\widetilde{\Pi}(x,p)=(\delta^{-1}p)\mathbb{E}(x\wedge M^0)-r\mathbb{E}(x\wedge\tau^0\wedge M^0)+c(x\wedge\tau^0)$. Because $\widetilde{\Pi}(x,p)$ is quasi-convex in x when $x\geq 0$, and it has roots $x=0$ and $x=\widetilde{s}(p)>0$, the integral $\widetilde{F}(x,p)=\int_0^x\widetilde{\Pi}(z,p)dz$ is decreasing in x when $0\leq x\leq\widetilde{s}(p)$, and then increasing in x when $x\geq\widetilde{s}(p)$. Thus, $\widetilde{s}(p)$ minimizes $\widetilde{F}(x,p)$ over all $x\geq 0$. In addition, clearly $\widetilde{\Pi}(x,p)$ is increasing in p, implying $\widetilde{F}(x,p)$ is supermodular in (x,p). By Theorem 2.8.2 in Topkis (1998), $\widetilde{s}(p)$ is positive and decreasing in p. Since $d_1(p)$ is also positive and decreasing in p, it follows that $s(p)=d_1(p)\widetilde{s}(p)$ is also decreasing in $p\in[P_{\min},P_{\max}]$. To sum up, we conclude the monotonicity of $s(p)$ when $p\leq\delta r$.

In particular, if M^0 follows the 2-point distribution in (EC.1), then $\tau^0=1-\kappa$ if $qr\leq c$ and $\tau^0=1+\kappa$ otherwise. It is straightforward to see $P_{\min}$ is decreasing in κ. Moreover, if $P_{\min}<p<P_{\max}$, then $\Pi(z,p)$ given by (EC.9) is linearly decreasing when $z\leq d_1(p)(1-\kappa)$, linear increasingly when $d_1(p)(1-\kappa)\leq z\leq d_1(p)(1+\kappa)$, and then remains to a constant when $z\geq d_1(p)(1+\kappa)$. By some basic calculations, one can verify that $s(p)=\frac{(1-\kappa)(\delta r-p)[\delta(r-c)-(1-q)p-\delta(qr-c)^+]}{\delta(1-\delta)[qp-\delta(qr-c)^+]}$, decreasing in κ. Thus, $s=s(p_m)$ is decreasing in κ. □

Proof of Proposition 9

To derive consumers' purchasing behavior, we first examine equations (EC.4) to (EC.6) closely. For given λ_1 and λ_2, we derive the range of v implied by each inequality.

- Inequality (EC.4): Note that it implies $\delta v-p-\frac{\phi}{\lambda_1}\geq[\lambda_2(v-r)-\phi]^+$. That is, $\delta v-p-\frac{\phi}{\lambda_1}\geq 0$ and $\delta v-p-\frac{\phi}{\lambda_1}\geq\lambda_2(v-r)-\phi$. Equivalently, $v\geq\frac{1}{\delta}[p+\frac{\phi}{\lambda_1}]$ and $(\delta-\lambda_2)v\geq p+\frac{\phi}{\lambda_1}-\phi-\lambda_2 r$.
- Inequality (EC.5): It is equivalent to $\lambda_2(v-r)-\phi\geq 0$ and $\lambda_2(v-r)-\phi\geq\lambda_1(\delta v-p)+(1-\lambda_1)[\lambda_2(v-r)-\phi]^+-\phi$. That is, $v\geq r+\frac{\phi}{\lambda_2}$ and $(\delta-\lambda_2)v\leq p+\frac{\phi}{\lambda_1}-\phi-\lambda_2 r$.
- Inequality (EC.6): It is equivalent to $0\geq\lambda_1(\delta v-p)+(1-\lambda_1)[\lambda_2(v-r)-\phi]^+-\phi$ and $0\geq\lambda_2(v-r)-\phi$. That is, $0\geq\lambda_1(\delta v-p)-\phi$ and $0\geq\lambda_2(v-r)-\phi$. Equivalently, $v\leq\frac{1}{\delta}[p+\frac{\phi}{\lambda_1}]$ and $v\leq r+\frac{\phi}{\lambda_2}$.

Note that $\frac{1}{\delta}[p+\frac{\phi}{\lambda_1}]\leq r+\frac{\phi}{\lambda_2}$ if and only if $p\leq\delta r+\phi(\frac{\delta}{\lambda_2}-\frac{1}{\lambda_1})$. Also, $\frac{1}{\delta-\lambda_2}(p+\frac{\phi}{\lambda_1}-\phi-\lambda_2 r)-\frac{1}{\delta}(p+\frac{\phi}{\lambda_1})=\frac{\lambda_2}{\delta(\delta-\lambda_2)}(p+\frac{\phi}{\lambda_1}-\delta r-\frac{\delta\phi}{\lambda_2})$ and $\frac{1}{\delta-\lambda_2}(p+\frac{\phi}{\lambda_1}-\phi-\lambda_2 r)-(r+\frac{\phi}{\lambda_2})=\frac{1}{\delta-\lambda_2}(p+\frac{\phi}{\lambda_1}-\delta r-\frac{\delta\phi}{\lambda_2})$. Now consider the following cases:

- $\lambda_2\leq\delta$ and $p\leq\delta r+\phi(\frac{\delta}{\lambda_2}-\frac{1}{\lambda_1})$: those with $v\geq\frac{1}{\delta}[p+\frac{\phi}{\lambda_1}]$ attempt to buy in the first phase, and those with $v\leq\frac{1}{\delta}[p+\frac{\phi}{\lambda_1}]$ do not make an attempt to buy in either phase. That is, $\alpha=1$ and $\rho=\frac{[1-(r+\frac{\phi}{\lambda_2})]^+}{1-\frac{1}{\delta}(p+\frac{\phi}{\lambda_1})}$.

- $\lambda_2 \le \delta$ and $p > \delta r + \phi(\frac{\delta}{\lambda_2} - \frac{1}{\lambda_1})$: those with $v \ge \frac{1}{\delta-\lambda_2}(p + \frac{\phi}{\lambda_1} - \phi - \lambda_2 r)$ attempt to buy in the first phase, those with $v \in [r + \frac{\phi}{\lambda_2}, \frac{1}{\delta-\lambda_2}(p + \frac{\phi}{\lambda_1} - \phi - \lambda_2 r)]$ attempt to buy in the second phase, and those with $v \le r + \frac{\phi}{\lambda_2}$ do not attempt to buy in either phase. That is, $\alpha = \frac{[1-\frac{1}{\delta-\lambda_2}(p+\frac{\phi}{\lambda_1}-\phi-\lambda_2 r)]^+}{1-(r+\frac{\phi}{\lambda_2})}$, and $\rho = 1$.
- $\lambda_2 > \delta$ and $p \le \delta r + \phi(\frac{\delta}{\lambda_2} - \frac{1}{\lambda_1})$: those with $v \in [\frac{1}{\delta}(p + \frac{\phi}{\lambda_1}), \frac{1}{\delta-\lambda_2}(p + \frac{\phi}{\lambda_1} - \phi - \lambda_2 r)]$ attempt to buy in the first phase, those with $v \ge \frac{1}{\delta-\lambda_2}(p + \frac{\phi}{\lambda_1} - \phi - \lambda_2 r)$ attempt to buy in the second phase, and those with $v \le \frac{1}{\delta}(p + \frac{\phi}{\lambda_1})$ do not attempt to buy in either phase. That is, $\alpha = \frac{1\wedge[\frac{1}{\delta-\lambda_2}(p+\frac{\phi}{\lambda_1}-\phi-\lambda_2 r)]-\frac{1}{\delta}(p+\frac{\phi}{\lambda_1})}{1-\frac{1}{\delta}(p+\frac{\phi}{\lambda_1})}$, and $\rho = \frac{\{[1\wedge(\frac{1}{\delta-\lambda_2}(p+\frac{\phi}{\lambda_1}-\phi-\lambda_2 r))]-(r+\frac{\phi}{\lambda_2})\}^+}{1\wedge[\frac{1}{\delta-\lambda_2}(p+\frac{\phi}{\lambda_1}-\phi-\lambda_2 r)]-\frac{1}{\delta}(p+\frac{\phi}{\lambda_1})}$.
- $\lambda_2 > \delta$ and $p > \delta r + \phi(\frac{\delta}{\lambda_2} - \frac{1}{\lambda_1})$: those with $v \ge r + \frac{\phi}{\lambda_2}$ attempt to buy in the second phase, and those with $v \le r + \frac{\phi}{\lambda_2}$ do not attempt to buy in either phase. Thus, $\alpha = 0$ and ρ is arbitrary.

$\square$

Proof of Proposition 10

Given the market demand M, parameters α, ρ related to inventory rationing, and cost parameters r and c, the single-period profit corresponding to optimal $y^* = \tau$ and a general $0 \le z \le x$ is given by

$$\begin{aligned}\pi(\tau, z) &= p\mathbb{E}(z \wedge \alpha M) - c(\beta\tau + \rho(\tau - z)^+) + r\mathbb{E}(\beta(\tau \wedge \alpha M) + \rho(\tau \wedge \alpha M - z)^+)\\ &= (\beta + \rho)[r\mathbb{E}(\tau \wedge \alpha M) - c\tau] + p\mathbb{E}(z \wedge \alpha M) - \rho\left[r\mathbb{E}(z \wedge \tau \wedge \alpha M) - c(z \wedge \tau)\right],\end{aligned}$$

where $\beta = \frac{1-\alpha}{\alpha}$ and τ maximizes $r\mathbb{E}(x \wedge \alpha M) - cx$. We are interested in the cutoff level s, which is defined as the (larger) root of $\pi(\tau, z) = 0$ or equivalently, $f(z) = 0$ for

$$f(z) = p\mathbb{E}(z \wedge \alpha M) - \rho\left[r\mathbb{E}(z \wedge \tau \wedge \alpha M) - c(z \wedge \tau)\right].$$

Suppose $\mathcal{P}(M = M_{\min}) = 1 - q$ and $\mathcal{P}(M = M_{\max}) = q$ for some $0 \le M_{\min} \le M_{\max}$ and $0 \le q \le 1$. Then $f(z)$ consists of three linear pieces. In particular,

$$f(z) = \begin{cases} [p - \rho(r - c)]z, & \text{if } z \le \alpha M_{\min} \\ \alpha p\mathbb{E}M - \rho[r\mathbb{E}(\tau \wedge \alpha M) - c\tau], & \text{if } z \ge \alpha M_{\max}. \end{cases}$$

and if $\alpha M_{\min} \le z \le \alpha M_{\max}$, then

$$\begin{aligned} f(z) &= \begin{cases} p\mathbb{E}(z \wedge \alpha M) - \rho(r - c)\alpha M_{\min} & \text{if } \tau = \alpha M_{\min} \\ (p - \rho r)\mathbb{E}(z \wedge \alpha M) + \rho c z, & \text{if } \tau = \alpha M_{\max} \end{cases} \\ &= \begin{cases} p[(1-q)\alpha M_{\min} + qz] - \rho(r - c)\alpha M_{\min} & \text{if } \ qr \le c \\ (p - \rho r)[(1-q)\alpha M_{\min} + qz] + \rho c z, & \text{if } \ qr \ge c \end{cases} \end{aligned}$$

To find the cutoff s, consider the following cases:

- $c \geq qr$: $\tau = \alpha M_{\min}$. If $p \geq \rho(r-c)$, $f(z)$ always increases in z, and thus $s=0$. If $\rho(r-c)\frac{M_{\min}}{\mathbb{E}(M)} < p < \rho(r-c)$, $f(z)$ decreases and then increases in z, attaining a positive value at $z = \alpha M_{\max}$. Hence, $s = \left[\frac{\rho(r-c)-p(1-q)}{qp}\right]\alpha M_{\min}$. If $p \leq \rho(r-c)\frac{M_{\min}}{\mathbb{E}(M)}$, $f(z)$ first decreases and then increases in z, attaining a nonpositive value at $z = \alpha M_{\max}$. Hence, $s = +\infty$. Since the initial state never exceeds $M_{\max}$, it suffices to set $s = M_{\max}$.
- $c < qr$: $\tau = \alpha M_{\max}$. If $p \geq \rho(r-c)$, $f(z)$ always increases in z, and thus $s=0$. If $\rho(r - c\frac{M_{\max}}{\mathbb{E}(M)}) < p < \rho(r-c)$, $f(z)$ decreases and then increases in z, attaining a positive value at $z = \alpha M_{\max}$. Hence, $s = \left[\frac{(\rho r-p)(1-q)}{\rho c - q(\rho r - p)}\right]\alpha M_{\min}$. If $p \leq \rho(r - c\frac{M_{\max}}{\mathbb{E}(M)})$, $f(z)$ first decreases and then may increase in z, attaining a nonpositive value at $z = \alpha M_{\max}$. Hence, $s = +\infty$. Since the initial state never exceeds $M_{\max}$, it suffices to set $s = M_{\max}$. □

Proof of Corollary 1

Suppose $c < qr$. $\tau = \alpha M_{\max}$. It is easy to see that $\lambda_2^* = 1$. By definition of λ_1^*:

$$\begin{aligned}\lambda_1^* = (1-q)\mathbb{E}_{X_0}&\left[\frac{X_0}{\alpha M_{\min}}\mathbb{I}_{s \leq X_0 \leq \alpha M_{\min}} + \mathbb{I}_{X_0 > s \vee (\alpha M_{\min})}\right] \\ &+ q\mathbb{E}_{X_0}\left[\frac{X_0}{\alpha M_{\max}}\mathbb{I}_{s \leq X_0 \leq \alpha M_{\max}} + \mathbb{I}_{X_0 > s \vee (\alpha M_{\max})}\right].\end{aligned}$$

By Proposition 10, consider the following three cases:

- If $p \geq \rho(r-c)$, then $s = 0$ and

$$\begin{aligned}\lambda_1^* =& (1-q)\mathbb{E}_{X_0}\left[\frac{X_0}{\alpha M_{\min}}\mathbb{I}_{0 \leq X_0 \leq \alpha M_{\min}} + \mathbb{I}_{X_0 > \alpha M_{\min}}\right] + q\mathbb{E}_{X_0}\left[\frac{X_0}{\alpha M_{\max}}\mathbb{I}_{0 \leq X_0 \leq \alpha M_{\max}} + \mathbb{I}_{X_0 > \alpha M_{\max}}\right] \\ =& (1-q)\left[\frac{\alpha M_{\min}}{2M_{\max}} + \frac{M_{\max} - \alpha M_{\min}}{M_{\max}}\right] + q\left(\frac{\alpha}{2} + 1 - \alpha\right) \\ =& (1-q)\left(1 - \frac{\alpha}{2}\frac{M_{\min}}{M_{\max}}\right) + q\left[1 - \frac{\alpha}{2}\right] = 1 - \frac{\alpha}{2}\frac{\mathbb{E}M}{M_{\max}}.\end{aligned}$$

- If $\rho\left[r - c\frac{M_{\max}}{\mathbb{E}(M)}\right] < p < \rho(r-c)$, then $s = \left[\frac{(\rho r-p)(1-q)}{\rho c - q(\rho r-p)}\right]\alpha M_{\min}$, $\alpha M_{\min} < s < \alpha M_{\max}$, and

$$\begin{aligned}\lambda_1^* =& (1-q)\mathbb{E}_{X_0}\left[\mathbb{I}_{X_0 > s}\right] + q\mathbb{E}_{X_0}\left[\frac{X_0}{\alpha M_{\max}}\mathbb{I}_{s \leq X_0 \leq \alpha M_{\max}} + \mathbb{I}_{X_0 > \alpha M_{\max}}\right] \\ =& (1-q)(1 - \frac{s}{M_{\max}}) + q\left[\frac{1}{2}(\alpha - \frac{s^2}{\alpha M_{\max}^2}) + 1 - \alpha\right] \\ =& 1 - \frac{s}{M_{\max}} - \frac{q}{2\alpha}\left(\alpha - \frac{s}{M_{\max}}\right)^2 \\ =& 1 - \left[\frac{(\rho r-p)(1-q)}{\rho c - q(\rho r-p)}\right]\alpha\frac{M_{\min}}{M_{\max}} - \frac{q\alpha}{2}\left(1 - \left[\frac{(\rho r-p)(1-q)}{\rho c - q(\rho r-p)}\right]\frac{M_{\min}}{M_{\max}}\right)^2.\end{aligned}$$

- If $p \leq \rho\left[r - c\frac{M_{\max}}{\mathbb{E}(M)}\right]$, then $s = M_{\max}$ and $\lambda_1^* = 0$

□

Proof of Proposition 11

For consumers' best response, we first note that, since $\lambda_2^*=1$ (from Corollary 1), $\left[\frac{1}{\delta}\left(p+\frac{\phi}{\lambda_1^*}\right)\right]\wedge\left(r+\frac{\phi}{\lambda_2^*}\right)<1$ is always satisfied, by Assumption (A.2). Hence, Proposition 9 applies. Furthermore, among the four cases in the proposition, only the case (iii) can be sustained in equilibrium. To see why, note that $\lambda_2^*=1>\delta$, implying that neither case (i) nor (ii) can be in equilibrium. Suppose case (iv) is in equilibrium, then $\alpha^*=0$ implies $\lambda_1^*=1$ (see points (a) and (b) in Corollary 1). Substituting $\lambda_1^*=1$ into the condition of case (iv), we have $p>\delta r-\phi(1-\delta)$, which violates Assumption (A.3). Substituting $\lambda_2=1$ to case (iii) in Proposition 9 and noting Assumption (A.2), we have: For $p\le\delta r+\phi(\delta-\frac{1}{\lambda_1})$,

$$\alpha^*(\lambda_1,1)=\frac{1\wedge\left[\frac{1}{1-\delta}(\phi+r-p-\frac{\phi}{\lambda_1})\right]-\frac{1}{\delta}(p+\frac{\phi}{\lambda_1})}{1-\frac{1}{\delta}(p+\frac{\phi}{\lambda_1})},$$

$$\rho^*(\lambda_1,1)=\frac{1\wedge\left[\frac{1}{1-\delta}(\phi+r-p-\frac{\phi}{\lambda_1})\right]-(r+\phi)}{1\wedge\left[\frac{1}{1-\delta}(\phi+r-p-\frac{\phi}{\lambda_1})\right]-\frac{1}{\delta}(p+\frac{\phi}{\lambda_1})}.$$

Since we focus on the case $\alpha^*(\lambda_1,1)<1$, we further assume $\phi+r+\delta-1-\frac{\phi}{\lambda_1}<p\le\delta r+\phi(\delta-\frac{1}{\lambda_1})$. In such a case,

$$\alpha^*(\lambda_1,1)=\frac{\frac{1}{1-\delta}(\phi+r-p-\frac{\phi}{\lambda_1})-\frac{1}{\delta}(p+\frac{\phi}{\lambda_1})}{1-\frac{1}{\delta}(p+\frac{\phi}{\lambda_1})}, \tag{EC.10}$$

$$\rho^*(\lambda_1,1)=\delta. \tag{EC.11}$$

For the firm's best response in Corollary 1, since we focus on the equilibria with $\lambda^*>0$, we only need to consider cases (a) and (b) in Corollary 1. That is, $\lambda_2^*=1$ and λ_1^* satisfies

$$\lambda_1=1-\alpha W_0(\rho) \tag{EC.12}$$

$$\text{and } W_0(\rho)=\begin{cases}\frac{\mathbb{E}(M)}{2M_{\max}} & \text{if } p\ge\rho(r-c)\\ \left[\frac{(\rho r-p)(1-q)}{\rho c-q(\rho r-p)}\frac{M_{\min}}{M_{\max}}+\frac{q}{2}\left(1-\frac{(\rho r-p)(1-q)}{\rho c-q(\rho r-p)}\frac{M_{\min}}{M_{\max}}\right)^2\right], & \text{if } \rho(r-c\frac{M_{\max}}{\mathbb{E}(M)})\le p\le\rho(r-c)\end{cases} \tag{EC.13}$$

In equilibrium, equations (EC.10) through (EC.13) hold simultaneously. That is, $\alpha=\alpha^*(\lambda_1,1)$ and $\rho=\rho^*(\lambda_1,1)$. Substituting equation (EC.10) to equation (EC.12) and letting $W=W_0(\delta)$, we have

$$\lambda_1=1-W\frac{\frac{1}{1-\delta}(\phi+r-p-\frac{\phi}{\lambda_1})-\frac{1}{\delta}(p+\frac{\phi}{\lambda_1})}{1-\frac{1}{\delta}(p+\frac{\phi}{\lambda_1})}$$

$$\Leftrightarrow g(\lambda_1)=(1-\delta)(\delta-p)\lambda_1^2+(W(r\delta-p+\delta\phi)+(\delta-1)(\delta-p+\phi))\lambda_1-(\phi W+\phi(\delta-1))=0$$

Since $\delta < 1$ (by Assumption A.1) and $p < \delta$ (by Assumptions A.2 and A.3), $g(\lambda_1)$ is strictly convex in λ_1. To sustain an equilibrium, $g(\lambda_1)$ must have a root satisfying $\phi + r + \delta - 1 - \frac{\phi}{\lambda_1} < p \leq \delta r + \phi(\delta - \frac{1}{\lambda_1})$, or $\phi + r + \delta - 1 - p < \frac{\phi}{\lambda_1} \leq \delta r + \phi\delta - p$. Note that Assumption (A.4) implies $\phi + r + \delta - 1 - p < \phi$. Hence, $\phi + r + \delta - 1 - p < \frac{\phi}{\lambda_1}$ is satisfied by any $\lambda \in (0,1]$. Meanwhile, Assumption (A.3) guarantees $\delta r + \phi\delta - p > 0$. Hence, the equilibrium condition becomes $\frac{\phi}{\delta(r+\phi)-p} < \lambda_1 \leq 1$. By Assumptions (A.2) and (A.3), $g(\frac{\phi}{\delta(r+\phi)-p}) = \delta\phi\,(1-\delta)\,\frac{1-(r+\phi)}{(r\delta-p+\delta\phi)^2}\,(p+\phi-r\delta-\delta\phi) < 0$ and $g(1) = -W\,(p+\phi-r\delta-\delta\phi) > 0$. Hence, there exists a unique solution λ_1^* to $g(\lambda_1) = 0$ satisfying $\frac{\phi}{\delta(r+\phi)-p} < \lambda_1 \leq 1$. Furthermore, $g'(\lambda_1) > 0$ at $\lambda_1 = \lambda_1^*$. Also note that

$$
\begin{aligned}
\frac{d}{d\phi}h(\lambda_1) &= \delta\lambda_1 - \lambda_1 - W - \delta + \delta\lambda_1 W + 1 \\
&= (1-\lambda_1)((1-W-\delta) - (1-\delta)W\frac{\lambda_1}{1-\lambda_1}) \\
&\leq (1-\lambda_1)((1-W-\delta) - (1-\delta)W\frac{1-W}{W}) \\
&= -\delta W\,(1-\lambda_1) < 0
\end{aligned}
$$

where the first inequality is due to the facts $\lambda_1 = 1 - \alpha W$ and $\alpha \in [0,1]$, implying $\frac{1-\lambda_1}{W} \leq 1$, i.e., $\frac{\lambda_1}{1-\lambda_1} \geq \frac{1-W}{W}$. Hence, λ_1^* increases in ϕ. By (EC.12) and the fact $W_0(\delta)$ is a constant independent of ϕ, α^* decreases in ϕ. This result, together with Proposition 10, implies that the cutoff s (weakly) decreases in ϕ. $\square$